EFFECTS OF PRESCRIBED BURNING ON AIR QUALITY IN THE SOUTHEASTERN U.S. AND IMPLICATIONS FOR PUBLIC HEALTH STUDIES

A Dissertation
Presented to
The Academic Faculty

by

Ran Huang

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
May 2019

EFFECTS OF PRESCRIBED BURNING ON AIR QUALITY IN THE SOUTHEASTERN U.S. AND IMPLICATIONS FOR PUBLIC HEALTH STUDIES

Approved by:

Dr. Armistead G. Russell, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Yongtao Hu
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. M. Talat Odman, Co-advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Jennifer Kaiser
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. James A. Mulholland
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Rodney J. Weber
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Date Approved: March 12, 2019

ACKNOWLEDGEMENTS

I would like to express my sincerest thanks to all the wonderful people I have met and who have supported me during my graduate studies at Georgia Tech. First and foremost, I extend my deepest thanks to my advisors Dr. Armistead Russell and Dr. Talat Odman, for their insightful and generous guidance and inspiration throughout my dissertation work. I could not have finished this work without their help. Also thanks to the Russell Group members, past and present, especially Dr. Cesunica E. Ivey and Dr. Xinxin Zhai. They all shared their tremendous kindness and help with me and made my life at Georgia Tech unforgettable. I also want to thank my thesis committee members, Dr. James Mulholland, Dr. Jennifer Kaiser, Dr. Rodney Weber, and Dr. Yongtao Hu for their precious time and constructive feedback. Special thanks to Dr. Hu, he has been very supportive when I had trouble running my model.

I greatly appreciate all the help and encouragement I received from my family and friends. Special thanks to my parents who provide incredible support and love to me both materially and spiritually. Thanks to my friend at Tech, Dr. Fei He, Dr. Xiaofei Zeng, Dr. Qian Zhang, and many others for the fun times we had. Thanks to my friends in California, Miss Yiwei Sun and Miss Xuelun Gong for all their encouragements. Thanks to my friends in China, Miss Di Chen, Mr. Wenjun Yang, Miss Dan, and many others, for sincerely listening to my chatter and providing absolute support, even with a jet lag. Last but not the least, I would like to seriously thank my medicine and calligraphy teachers, Mr. Jianhua Wang and Miss Kun Qian. I really appreciate the skills they taught me. I had some tough times during my studies, but those skills helped me calm down and keep me going.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AAI	absorbing aerosol index
AOD	aerosol optical depth
AQI	Air Quality Index
AVHRR	Advanced Very High Resolution Radiometer
BAECV	Burned Area Essential Climate Variable
BAM	β-attenuation monitors
BB	biomass burning
(Blended-)BBEP	Blended Polar Geo Biomass Burning Emissions Product
BenMAP-CE	Environmental Benefits Mapping and Analysis Program – Community Edition
CF	correction factor
CFPP	coal-fired power plants
CMAQ	Community Multiscale Air Quality
CMB	chemical mass balance
CO	Carbon Monoxide
CONUS	Contiguous United States
CSN	Chemical Speciation Network
CTM	chemical transport model
DASI	Dispersive Apportionment of Source Impacts
DCHS	Dougherty County High School
DDM	Decoupled Direct Method
DF	data fusion
dNBR	difference normalized burn ratio

DV	diesel vehicle
EC	Elemental Carbon
EGU	Electric Generating Units
EPA	Environmental Protection Agency
EPD	Environmental Protection Division
EPM	Emissions Production Model
ER	Emergency Room
EST	Eastern Standard Time
ETM+	Enhanced Thematic Mapper plus
FCCS	Fuel Characteristic Classification System
FEMs	Federal Equivalent Methods
FEPS	Fire Emissions Production Simulator
FFS	Florida Forest Service
FIMMA	Fire Identification, Mapping, and Modeling Algorithm
FL	Florida
FOFEM	First Order Fire Effects Model
GA	Georgia
GBS	Global Burned Surfaces
GFC	Georgia Forestry Commission
GFED	Global Fire Emissions Database
GV	gasoline vehicle
GWEM	Global Wildland Fire Emission Model
HANDS	Hotspot and NDVI Differencing Synergy
HMS	Hazard Mapping System
HYSPLIT	Hybrid Single Particle Lagrangian Integrated Trajectory

IMPROVE	Interagency Monitoring of Protected Visual Environments
IMSI	Integrated Mobile Source Indicator
IQR	interquartile range
JST	Jefferson Street
LCHS	Lee County High School
LOO	leave-one-monitor-out
MODIS	Moderate Resolution Imaging Spectroradiometer
mRMSD	maximum daily root-mean-squared-deviation
NAAQS	national ambient air quality standards
NASA	National Aeronautics and Space Administration
NC	North Carolina
NEI	National Emission Inventory
NH_4^+	Ammonium
NME	Normalized Mean Error
NO_2	Nitrogen Dioxide
NO_3^-	Nitrate
NOAA	National Oceanic and Atmospheric Administration
NOx	Nitrogen Oxides
NRMSE	Normalized Root Mean Square Error
NWS	National Weather Service
OBS	observation
OC	Organic Carbon
OSAT	ozone source apportionment
prescribed burning	prescribed burning
PM	particulate matter

$PM_{2.5}$	particulate matter with an aerodynamic diameter less than 2.5 μm
PSAT	particulate matter source apportionment
RH	relative humidity
RMSE	Root Mean Square Error
SIM	simulation
SLAMS	State and Local Air Monitoring Stations
SMOKE	Sparse Matrix Operator Kernel Emissions model
SO_4^{2-}	Sulfate
SOA	secondary organic aerosol
TEOM	tapered element oscillating microbalance
U.S.	the United States
USGS	U.S. Geological Survey
VISTAS	Visibility Improvement – State and Tribal Association of the Southeast
WCHS	Worth County High School
WF-ABBA	WildFire Automated Biomass Burning Algorithm
WH	withholding
WHO	World Health Organization
WRF	Weather Research and Forecasting model

SUMMARY

Exposure to air pollution has been associated with increased morbidity and premature mortality, indicating that sustained reductions in pollution exposure could lead to improved health and increased life expectancy. Biomass burning is an important global source of gases and aerosols, e.g., carbon monoxide, carbon dioxide, $PM_{2.5}$ (particulate matter with an aerodynamic diameter less than 2.5 μm) and black carbon. These products generally referred to as "smoke" can reduce visibility and have adverse health effects. Prescribed burning, a type of biomass burning, is a land management practice used in the U.S. to reduce wildfire risk and maintain healthy ecosystems. This dissertation is a presentation of research results quantifying the impact of prescribed burning on air quality and human health in the southeastern U.S., the most active prescribed burning area in the U.S.

Considering the potential impacts of prescribed burning, the estimation of those emissions is crucial. The emissions estimates from the National Emission Inventory are based on the burned areas reported by the states, which may be subject to significant uncertainty since not all prescribed burns have reliable records. Satellite-derived products could be used as a substitute tool to provide burned area data. In order to evaluate burned areas from satellite-derived products and assess whether they can be used in prescribed fire burned area estimation, we conducted a comparison between prescribed burning permit records and satellite-derived burned areas for Georgia and Florida on the first four months of 2015 and 2016, which is the most active burn season in those two states, with two satellite-derived products: Blended Polar Geo Biomass Burning Emissions Product and

Global Fire Emissions Database. The comparison results indicate that both satellite-derived data products underestimate the burned areas compared to permit recorded data. Overall, current satellite-derived products have limitations in estimating the burned areas of small fires and still need improvements.

Another need is to split the combined prescribed fire impact derived from chemical transport models (CTMs) into individual fire impacts. A novel source apportionment method (Dispersive Apportionment of Source Impacts) has been developed for this by using concentration fields derived from dispersion modeling. Individual burn impacts obtained in this manner could help local land and air quality managers decide which burns should be allowed or restricted based on their impacts on air quality and public health in areas of concern.

The feasibility of applying low-cost PM sensors for the detection of fire impacts has been evaluated. The observations from the low-cost PM sensor were compared with the nearby reference instruments and simulation from a CTM. It was found that low-cost PM sensors can provide spatial information that is missed by a sparse regulatory monitoring network and, in combination with CTM simulations, they can be used in preparing high accuracy exposure fields needed for health assessments. Data fusion is a method that integrates observations from sensors/monitors with simulations from CTM to better estimate ground-level air pollutant concentrations. This method has been applied to North Carolina from 2006 to 2008 to support the University of North Carolina at Chapel Hill's health analysis of coronary heart disease patients by developing spatiotemporal exposure fields for $PM_{2.5}$ mass, five PM species, and three gases at a spatial resolution of 12 km. An inter-comparison was also performed of total $PM_{2.5}$ mass between the fields

using data fusion and two methods that included satellite aerosol optical depth (AOD) data. The results show that the data fusion method performs best among the three methods in this NC application. It has also been utilized to generate exposure fields to smoke from the prescribed fire. These fields have been input to a health impact function for asthma-related Emergency Room Visits to find the health impact due to the prescribed fire in Georgia during the burning season from 2015 to 2018. The spatial and temporal variation of health impact from prescribe burning illustrate the importance of distinguishing seasons and areas when looking at the relationship between pollutants exposure from prescribed fire and its health effects. Atlanta area has the largest health impact from prescribed burning with most population and moderate-level of fire impact.

Overall, the methods and results presented in this dissertation improve the understanding of the impact of prescribed burning on air quality and human health. The data generated would also benefit future health epidemiological studies. The work presented could be useful to scientists and policy makers interested in prescribed fire and its impact on air quality and public health.

CHAPTER 1. INTRODUCTION

1.1 Background and Motivation

The World Health Organization International Agency for Research on Cancer classifies that outdoor air pollution has been classified as carcinogenic to humans by WHO. Particulate matter, a major component of air pollution, is associated with increased incidences of cardiovascular disease (Brook et al. 2004). Over 4 million deaths (8% of total global mortality) were caused by exposure to outdoor $PM_{2.5}$ in 2015 (Forouzanfar et al. 2016). In the U.S., $PM_{2.5}$ is the environmental risk factor with the largest health burden and the 6^{th} largest mortality risk overall (Cohen et al. 2017).

Biomass burning, burning of land covering vegetation, is one of the most important global sources of $PM_{2.5}$. Exposure to smoke from biomass burning has been associated with adverse health effects (Reid et al. 2016). Many epidemiological studies have shown the associations between short-term $PM_{2.5}$ exposure from fires and health endpoints, like mortality (Faustini et al. 2015; Linares et al. 2015), respiratory effects (Dohrenwend et al. 2013; Johnston et al. 2014), and cardiovascular effects (Rappold et al. 2011; Yao et al. 2016).

Prescribed burning, a type of biomass burning, is a land management tool used to improve native vegetation and wildlife habitat, control insects and disease, and reduce wildfire risk. In the U.S., prescribed burning is very popular but its practice entails air pollution concerns. The U.S. Environmental Protection Agency (EPA) 2014 National Emission Inventory (NEI) reported that 14.8% of $PM_{2.5}$ emissions in the U.S. were

attributable to prescribed burning and 35% of $PM_{2.5}$ emissions from prescribed burning originated from the southeastern U.S. (US EPA 2014). Prescribed burning emissions remains as one of the most prominent sources of $PM_{2.5}$ in the southeastern U.S. (30% from 2014 NEI). Accurate estimation of fire emissions is the first step to modeling the impact of those fires on air quality and human health. Emissions estimates in NEI and other emission inventories are based on the burned areas reported by the states, which may be subject to large uncertainty since not all prescribed burns have reliable records. For example, in some states, no permit is required to conduct a burn. Some states keep records of the burns on state and private lands, but federal lands are not obligated to report burns to the state. Lacking permit data is a problem that needs to be solved. Also, the permit records themselves have uncertainties since they contain values for areas intended to be burned and not the actual burned areas.

A crucial problem encountered by prescribed burn managers is the conflict between the increased demand for burning and the desire for better air quality for health. As wildfire risk increases with changing the climate (Liu et al. 2010), reliance on prescribed burning will increase. Meanwhile, strict controls are curbing emissions from other sources of air pollution. All these factors are increasing the role of prescribed burning emissions in national air pollution.

Several programs through the U.S. are encouraging participation in forest restoration, such as longleaf pine habitat restoration program. With the management of those new forests comes the need to burn more than before. Land and air quality managers need to know how much burning they can allow per day so that air pollution levels will not exceed the national ambient air quality standards (NAAQS) and will not affect local

people's health. A better understanding of the contributions of prescribed burning to air pollution, climate change, and human health is important, especially to the people who are affected by prescribed burning directly. The land management community also needs a path to improve decision making to cope with the possibility of controls.

The most important concern associated with prescribed burning in the southeastern U.S. and intensifying burning activity in the future is the potential to increase human exposure to air pollution and the associated health impacts. Rappold et al. (2017) found that over 40% of Americans are estimated to live in areas with a moderate or high contribution of wildland fires to ambient $PM_{2.5}$ concentrations. The southeastern U.S. houses some of the most vulnerable communities in the nation, and this region is more likely to experience high and frequent smoke exposure in comparison to the other parts of the country.

1.2 Research Objectives

The overall goal of this study was to quantify the effect of prescribed burning on air quality and human health. The specific research objectives are:

1. To evaluate the current satellite-derived burned area by comparing with prescribed burning permits data.
2. To quantify the prescribed burning impact on air quality using low-cost sensor measurements and chemical transport model.
3. To quantify the prescribed burning impact on human health based on pollutant exposure fields using data fusion method.

1.3 Thesis Structure

This dissertation highlights the major impacts from prescribed burning on air quality and public health in the southeastern U.S. The studies presented in the chapters are briefly summarized below.

Chapter 2 evaluates current satellite-derived products and assesses whether they can be used in prescribed burning area estimation. An inter-comparison has been conducted between prescribed burning permit records and two satellite-derived products, Blended Polar Geo Biomass Burning Emissions Product (BBEP) and Global Fire Emissions Database (GFED4s), in Georgia and Florida during the burn season for 2015 and 2016.

Considering that policy makers and local air quality managers are often more interested in the impacts of a single prescribed burn when deciding if it should be allowed or if it will have a major impact on people or areas of concern (e.g., highways, airports), in Chapter 3, we introduce a novel source apportionment method (Dispersive Apportionment of Source Impacts (DASI)) and apply it to split the combined prescribed burning impact from chemical transport model (Community Multiscale Air Quality (CMAQ) (Byun and Schere 2006a)) using the decoupled direct method (Napelenok et al. 2006) (DDM) into individual burn impact by using a dispersion model (Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) (Stein et al. 2015)). We applied the method on clustered prescribed burns, including three fires with the same burned area but different emissions due to different fuel loadings and consumptions. One of the most important contributions of this new method is that not only it can be used to split combined fire impact, but it could also be applied to other emission sources as a source apportionment method.

Chapter 4 evaluates the feasibility of using a low-cost PM sensor as a supplemental measurement of $PM_{2.5}$ concentrations in southwestern Georgia, one of the most active

prescribed burning areas in the southeastern U.S. Three low-cost sensors (Plantower PMS 3003) were initially deployed at three high schools (Dougherty, Lee, and Worth County High Schools (DCHS, LCHS, WCHS)) to measure the local $PM_{2.5}$ concentrations starting in May 16, 2017 and a fourth low-cost sensor was placed next to the state monitoring site on March 14, 2018. Different calibration methods have also been evaluated. CMAQ was also used to simulate the contribution of prescribed burning on $PM_{2.5}$ concentration. We also compared the simulation with the observation obtained from both the low-cost sensors and the reference monitor. A recommendation has been presented that suggests fusing model simulations with observations from a dense network of low-cost sensors to provide accurate exposure fields from the smoke.

The fusion method mentioned above is called data fusion, a new method that combining monitor observations and simulated data from a chemical transport model to obtain spatiotemporal pollutant fields. There is a trend in the air pollution community to use combinations of modeled and observed air quality data to estimate air pollutant concentration fields for use in exposure estimation. In Chapter 5, we applied this method to North Carolina from 2006 to 2008 for $PM_{2.5}$ total mass, OC, EC, SO_4^{2-}, NO_3^-, NH_4^+, CO, NO_x, and NO_2. These resulting fields capture the spatiotemporal information provided by the air quality model, as well as the finer temporal scale variations from the pollutant observations and decrease model biases. Several data withholding methods are then conducted to evaluate the data fusion method. We also compared the $PM_{2.5}$ fields from data fusion with the other two methods that use satellite aerosol optical depth (AOD) data. We highlight the major advantage of CTM-based data fusion methods over methods relying

mostly on AOD is that it provides speciated $PM_{2.5}$ and gaseous pollutant fields which are important to epidemiological studies.

Given the influence fire can exert on $PM_{2.5}$ levels, assessing the impact of prescribed burning on public health is a major research need. In Chapter 6, we explored the health impact of prescribed burning in Georgia for the first four months from 2015 to 2018. A typical health impact function was utilized to calculate the health impact with Emergency Room (ER) visits due to asthma as the health endpoint. National health incidence rate for 2013 was used. Population data were extracted from the Environmental Benefits Mapping and Analysis Program – Community Edition (BenMAP-CE) (Sacks et al. 2018). The adjusted prescribed burning impact on $PM_{2.5}$ comes from CMAQ-DDM after applying data fusion method. We provide some recommendations to the epidemiological study according to the findings in this study. For instance, there are specific days and areas to focus on to investigate the relationship between health effects and prescribed burning. Also, the temporal and spatial variations of prescribed burning impact on human health need to be further studied.

Chapter 7 concludes with the major findings from the above studies. Plans for future research to further improve the understanding of the impacts of prescribed burning on air quality and human health are also presented.

CHAPTER 2. BURNED AREA COMPARISONS BETWEEN PRESCRIBED BURNING PERMITS IN SOUTHEASTERN USA AND TWO SATELLITE-DERIVED PRODUCTS

As published in *Journal of Geophysical Research: Atmospheres*

Abstract

Prescribed burning is one of the most prominent sources of $PM_{2.5}$ (particulate matter with an aerodynamic diameter less than 2.5 μm) in the southeastern US. The prescribed burning emissions estimates may have significant uncertainty because they are based on the burned areas reported to the state agencies when burners apply for burn permits. When no permit records are available, satellite-derived products could be used as a substitute tool to provide burned area data. In order to evaluate burned areas from satellite-derived products, we conducted a comparison between prescribed burning permit records and two satellite-derived products, Blended Polar Geo Biomass Burning Emissions Product (BBEP) and Global Fire Emissions Database (GFED4s), in Georgia and Florida. The comparison results indicate that both satellite-derived products underestimate seriously the burned areas compared to permit record data. They can capture a cluster of fires better than isolated fires but may misinterpret those small fires together as one big fire. Overall, current satellite-derived products have limitations in estimating the burned areas of small fires and still need improvements.

2.1 Introduction

Biomass burning (BB) is the burning of land covering vegetation for a wide range of purposes ranging from land clearing to restoring nutrients to the soil. Emissions from BB could interact with the atmosphere and climate systems, change carbon balances and atmospheric chemistry components, and affect clouds and precipitation (Liu 2005), permafrost structure, and surface albedo (Natarajan et al., 2012; Randerson et al., 2006; Sokolik et al., 2010). BB is also an important global source of aerosols such as $PM_{2.5}$ (particulate matter with an aerodynamic diameter less than $_{2.5}$ μm) and black carbon (Andreae and Merlet 2001), which can affect human health, and smoke, which can reduce visibility and paralyze highway transportation. Considering the potential impacts of BB, the estimation of those emissions is crucial. Typically, emissions are calculated by the amount of burned area, biomass present in the ecosystem, efficiency of combustion and pollutants emitted per unit mass of fuel consumed, a.k.a. emission factors (Seiler and Crutzen 1980). Here, we will focus on the burned area.

Prescribed burning is a type of BB employed as a land management tool since the 1930's to improve native vegetation and wildlife habitat, control insects and disease, and reduce wildfire risk in the U.S. (Leopold 1987). The U.S. Environmental Protection Agency (EPA) 2011 and 2014 National Emission Inventories (NEIs) reported that 14.8% (2014) and 15.1% (2011) of $PM_{2.5}$ emissions in the U.S. are attributable to prescribed burning and 35% of $PM_{2.5}$ emissions from prescribed burning originate from the southeastern U.S. (U.S. EPA, 2011, 2014). Prescribed burning is one of the most prominent sources of $PM_{2.5}$ emissions in the southeastern U.S. (20% in 2011 and 30% in 2014), and will become an increasing source as other sources are controlled. According to 2014 NEI, nearly 50% of all fire-related $PM_{2.5}$ emissions are from prescribed fires while over 75% of

those emissions come from prescribed fires in the southeastern states. These emissions estimates are based on the burned areas reported by the states, which may be subject to large uncertainty since not all states have reliable prescribed burn records. In some states, a permit is not necessary to conduct a burn. Some states keep records of the burns on state-owned and private lands but federal land managers are not obligated to report burns to the state. Also, the permit records themselves have uncertainties since they contain values for areas intended to be burned, not actual burned areas. Finding another, more reliable way to estimate those prescribed burned areas is important for accurately estimating the emissions of prescribed burning and evaluating its impacts to air quality.

Several satellite-derived products such as HANDS (Hotspot and NDVI Differencing Synergy) (Fraser et al., 2000a, 2000b), GBA2000 (Tansey et al., 2004), GWEM (Hoelzemann 2004), Global Fire Emissions Database 3 (GFED3) (Giglio et al., 2010), MCD45A1 (Roy et al., 2008), L3JRC (Tansey et al., 2008a, 2008b), GLOBSCAR (Simon et al., 2004), Global Burned Surfaces (GBS) (Carmona-Moreno et al. 2005), Global Fire Emissions Database (GFED4s) (Giglio et al., 2013) and Blended Polar Geo Biomass Burning Emissions Product (Blended-BBEP or BBEP in this paper) (Zhang and Kondragunta 2008) have been developed for estimating the burned area of BB. Only GFED4s and BBEP continue to be available publicly at this time. Several evaluations (Boschetti et al., 2004; Chuvieco et al., 2016; Hoelzemann, 2004; Kukavskaya et al., 2013; Li et al., 2000; Li et al., 2003; Randerson et al., 2012; Zhu et al., 2017) have been conducted to examine the uncertainties of those products. Although these products are in good agreement with emission inventories reported by different countries in most regions around the globe, there are major disagreements in terms of burned area estimates, which could be

caused by spatial, temporal and spectral disparities, distinct algorithms and instrument drift from different satellite products (Boschetti et al. 2004; Kukavskaya et al. 2013). For example, GWEM emissions are biased low in most regions compared to the ATSR-scaled MOZART inventory (Schultz 2002) but the uncertainty is the largest in regions where small fires dominate (Hoelzemann 2004). Those small fires may have a large impact on global BB carbon emissions (Randerson et al. 2012). Zhu et al. (2017) found large underestimation in croplands by comparing burned area from Moderate Resolution Imaging Spectroradiometer (MODIS) and other satellite products with higher resolution. Zeng et al. (2016) compared fire counts from MODIS with Visibility Improvement – State and Tribal Association of the Southeast (VISTAS) fire inventory and reported improved model performance with the MODIS-updated fire emission inventory. Hu et al. (2016) performed a comparison between Hazard Mapping System (HMS) fire product with permit records based on fire detection rate as a function of burned area. Up to now, uncertainty in satellite burned area estimates has been estimated only through comparison of different satellite products; there has been no comparison of satellite burned area estimates with ground-based burned area measurements or estimates.

In prescribed burning, fires are usually small to keep them under control. Satellite-derived products may have large uncertainty when used in estimating burned area of those small fires (Hoelzemann 2004; Randerson et al. 2012; Kukavskaya et al. 2013; Mouillot et al. 2014). This is especially true for current operational geostationary satellites whose imagers have a spatial footprint of 4 km in nadir view and as large as 8 – 12 km at the edge of the scan. This can lead to missed detections if fire temperature is not too high compared to surrounding pixels as the algorithm uses surrounding pixels to obtain background

temperature (Schroeder et al. 2010). In order to evaluate those satellite-derived products and assess whether they can be used in prescribed burning burned area estimation, we conducted a comparison between prescribed burning permit records and satellite-derived burned area. This paper will focus on the comparison of prescribed burning permit records in Georgia and Florida in the first four months of 2015 and 2016, which is the most active burn season in those two states, with two satellite-derived products: BBEP and GFED4s. The goal is to assess the uncertainty of those two products and to determine whether they can be used in follow-up research to forecast the prescribed burning impact on air quality.

2.2 Data

2.2.1 Burn permit data

2.2.1.1 Georgia prescribed burn permit record data

Georgia is one of the most active states in applying prescribed burning in the U.S. The Georgia Forestry Commission (GFC) is responsible for prescribed burning services in the state. It is necessary to obtain a burn permit from the local GFC office before burning woods, lands, marshes or other flammable vegetation. Burn permits contain the contact information of the landowner, county of burn, location of the burn, acres to be burned, the start and end times of the burn and the name/phone number of the person to contact during the burn in case additional information is needed according to Georgia Prescribed Burning Act (GA Code Ann. 12-6-145 – 12-6-149). The prescribed burning season for Georgia is from 1 October through 30 April. A burn ban goes into effect in 54 counties during the ozone season (1 May – 30 September). We obtained permit record data from GFC for the years 2015 and 2016. Considering the ozone season restriction and the relatively small

number of burns during October – December, we focused our analysis on the first four months of each year. Some permits have latitude/longitude information of the burn but most of them only have an address, which may be in a non-standard format that is difficult to be geo-referenced.

2.2.1.2 Florida open burn authorization record data

Florida, another high burn activity state in the U.S., issues 120,000 authorizations allowing landowners and agencies to prescribe burn an average of over 2 million acres (800,000 hectares) each year. Florida Forest Service (FFS), a division of Florida Department of Agriculture and Consumer Services, is in charge of managing prescribed fire in Florida. Landowners need to contact the local FFS office with their customer number and provide the location and size of their burn to get an authorization. A smoke plume model is executed before approving the burn request to make sure there are no potential problems with the smoke from the burn. The dominant burn types in Florida are agricultural, land clearing and silvicultural burns. We obtained open burn authorization (permit) data from FFS for the years 2015 and 2016 and conducted an analysis of the first four months of each year. Each open burn authorization in Florida has latitude and longitude information, which are either provided by a GPS device or from a digitized map during the authorization process.

2.2.1.3 Uncertainty of permit records

The permit records are considered to be accurate by the permit issuing agencies based on their experience and anecdotal evidence but there is no scientific research to support this hypothesis. One of the strongest support for their hypothesis is that the permits

are issued on the day of the burn and, when a landowner calls for a permit, there is a strong intention to burn. However, weather may still play a role in whether a burn is conducted or not on that day after the permit is issued. If the weather onsite is not conducive to a successful burn, permit holders may choose not to exercise their right to burn.

Another source of inaccuracy may be the recorded size of the burn. The burned areas in permit records are not necessarily equal to the actual burned areas; they are simply estimates of areas planned to be burned. Even if the burn is conducted on the day the permit was issued for, the actual burned area may differ from the permitted area. Considering the difference between the fire weather forecast and actual field conditions during the burn, landowners may burn more or less land than they reported at the time of permit application. For example, landowners usually ask for a permit for the total area of their land but drainage areas are not burned and some sensitive ecosystem areas may be avoided.

Neither Georgia nor Florida follows up with the burners to find out the actual burned areas. We conducted a limited survey in Georgia to evaluate how well the burned area values in permit records represent the actual burned areas. We called a small subset of the landowners and asked them to report any differences between the actual burned areas and the areas on the permit records. Because of the difficulty many landowners had in retrieving older records for, we surveyed burns from 2016 and 2017. Our initial selection of the burners aimed to be representative of the sizes and geographic locations of the burns throughout the state but the respondents did not necessarily match these profiles.

2.2.2 Satellite-derived data

2.2.2.1 Blended Polar Geo Biomass Burning Emissions Product (Blended-BBEP)

The Blended-BBEP is a continuous product for North America of BB emissions with active fire data that are detected using WildFire Automated Biomass Burning Algorithm (WF_ABBA) from GOES (Prins and Menzel 1992; Prins and Menzel 1994), Fire Identification, Mapping, and Modeling Algorithm (FIMMA) from the Advanced Very High Resolution Radiometer (AVHRR) (Li et al. 2000a; He and Li 2012), and an enhanced contextual fire detection algorithm from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Giglio et al., 2003; Giglio et al., 2016). The Blended-BBEP is produced by blending, every six hours, fires detected from GOES-East, GOES-West, MODIS on both the NASA Terra and Aqua satellites, and AVHRR on NOAA-15/17/18. The outputs include burned area and emissions of the following species: $PM_{2.5}$, CO, CH_4, CO_2, total non-methane hydrocarbon, NH_3, N_2O, NO_X, and SO_2. The burned area is simulated using active fire observations from MODIS, AVHRR, and GOES for each GOES fire pixel (Zhang et al., 2011). Instantaneous fire size obtained from 30-minute GOES observation using WildFire Automated Biomass Burning Algorithm (WF-ABBA) is found to be an accurate representation of burned area for that time interval. The determination of fire size happens for approximately 20-30% of all WF-ABBA detected fires. For fires that are detected but not determined in size by the WF-ABBA algorithm, the size is simulated using climatological diurnal variation in fire size specific to biomass type where the fire is observed. If fire hot spots are detected only by polar-orbiting satellites, fire size is determined using a conversion factor that was derived by regressing fire hotpots with the burn scars detected from post-fire Landsat ETM+ (Enhanced Thematic Mapper plus) imagery. We downloaded the 2015 and 2016 BBEP data from the BBEP website (http://satepsanone.nesdis.noaa.gov/pub/FIRE/BBEP-geo/PREVIOUS_DAYS/).

2.2.2.2 Global Fire Emissions Database (GFED4s)

Global Fire Emissions Database provides monthly burned area and fire emissions by combining satellite information on fire activity and vegetation productivity. The current version, Version 4, has a spatial resolution of 0.25 degrees. GFED4s is the product including small fires, combining 1-km thermal anomalies (active fires) from Terra and Aqua and 500 m burned area observations from MODIS daily composition thermal anomaly/fire products (MOD14A1 and MYD14A1) (Randerson et al. 2012). Small fire burned area is estimated by computing the difference normalized burn ratio (dNBR) for these two sets (1-km and 500m) of active fires and then combining these observations with other information such as efficacy of the burned area detection algorithm, the frequency of satellite overpasses and the rate of movement of the fire front (Randerson et al. 2012). We obtained the monthly GFED4s data from Dr. James Randerson's group at University of California, Irvine.

2.2.2.3 Burned Area Essential Climate Variable (BAECV)

To test the effect of finer spatial resolution, we used BAECV as an additional product. BAECV is the burned area product from Landsat satellite developed by the U.S. Geological Survey (USGS) (Hawbaker et al. 2017). Landsat has a repeat cycle of 16 days and BAECV only has annual data over CONUS with 30m×30m resolution. BAECV uses a gradient boosted regression model to estimate the probability that pixel had burned, followed by a thresholding process to generate a binary burned or unburned classification. We downloaded the 2015 BAECV data from the USGS website (https://rmgsc.cr.usgs.gov/outgoing/baecv/BAECV_CONUS_v1_2017/).

2.2.3 Correlation Analyses

We conducted correlation analyses using Pearson and Spearman correction coefficients. Using daily state total burned areas from Georgia's permit records and BBEP, we performed a log-transformation of the data to test if Pearson correlation coefficient is suitable for our analysis. Since BBEP data may have zero burned area on some days, which, when log-transformed, would lead to negative infinity, we added 30 acres (which is approximately equal to BBEP's detection limit) to both permit-record and BBEP burned areas. Then we took the common logarithms ($\log_{10}$ (x)) of both permit and BBEP data and plotted their correlations. Finally, we calculated the residuals of both the untransformed and transformed data sets as $y_{obs} - y_{mod}$ where $y_{mod} = x \times \text{Slope} + \text{Intercept}$, y_{obs} is the BBEP data and x is the permit data before or after the log transformations.

2.3 Results and Discussions

Prescribed burning includes controlled fires conducted for the maintenance and protection of commercial timber stands, land clearing, agriculture, reduction of vegetative fuels for wildfire prevention, and management of fire-dependent ecosystems. Here, we considered all burn permits/authorizations with records of burned areas.

2.3.1 Burn permit survey

The target sample size of our phone call survey was approximately 10% of the total area that got burned during the calendar year. There was a total of 96 respondents in our survey. Many respondents are from professional companies that have complete records of the burns they conducted. To those that did not have detailed records of their burns, we

asked to report the total acreage of their land minus the known protected areas to get the approximate area they burned. The result shows good agreement between permitted and actual burned area with r^2 equal to 0.64 and 0.84 for 2016 and 2017, respectively (Figure 2-1). Each point on the plot represents the totals for one burner. The figures include 371 small (less than 60 acres), 228 medium (60 to 134 acres) and 215 large (larger than 135 acres) burn permits for 2017, and 346 small, 212 medium and 185 large permits for 2016. The survey results suggest larger uncertainty in 2016 permit records than 2017 since the regression line deviates more from the 1:1 line with a slope of 0.78 versus 0.83 and an intercept of 85 versus 46. This may be simply due to better remembrance of the more recent year's burns.

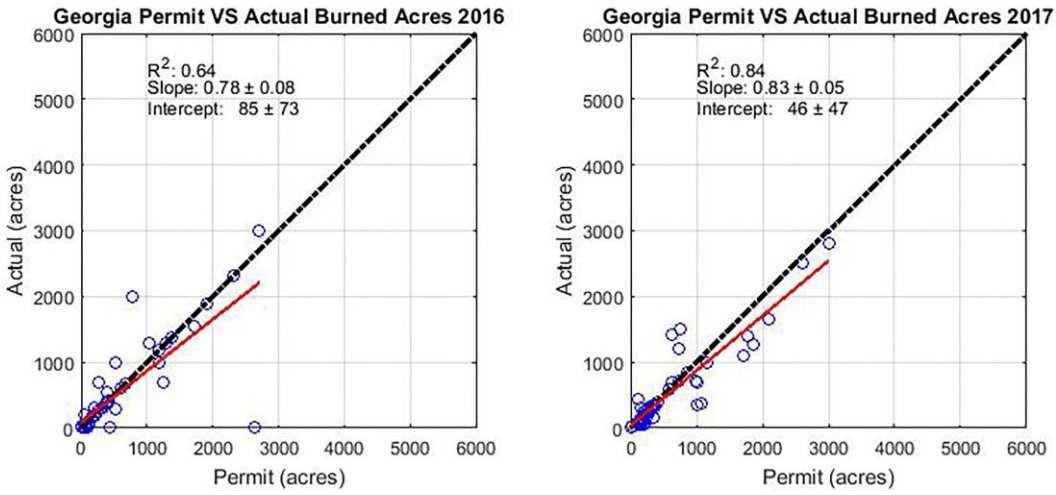

Figure 2-1 The results of a phone call survey comparing permitted and actual burned areas in Georgia for 2016 and 2017

The burned area data came from permit/authorization records. Actual burned areas are not tracked by the authorizing agencies (GFC and FFS) and the landowners are not required to call back and confirm or correct the burned areas. Our phone survey revealed that unburned areas within the plots such as drainages, deer camps and structures are not always excluded. However, our survey also found that this constitutes a small part of the

uncertainty. Unsuccessful burns remain in the records as permitted. GFC which has evaluated this issue before, postulates that since almost all of their permits were issued on the day of the burn, after the landowner reviewed the fire weather forecast and assessed the fuel conditions, the landowner's estimate is reliable. However, our survey discovered that a few burns were called off after the attempts of ignition, because it became obvious that the objectives would not be achieved. Considering all these factors, we placed the uncertainty of the burned area in the permit records at 20%.

2.3.2 Comparison of state, district and county total burned areas

Almost 50% of the prescribed burns in Georgia have areas smaller than 5 acres according to the permit records while, according to BBEP, nearly 80% of the fires have sizes ranging between 25 to 50 acres (Table A - 1). This inconsistency of the dominant fire size between the two datasets may be due to the satellite detection limit. In Florida, about 25% of the authorized burns have sizes between 25 to 50 acres, with another 20% ranging between 50 to 100 acres (Table A - 2). However, more than 45% of the fires BBEP captures are in the 25 to 50 acres size range, with less than 10% between 50 to 100 acres. BBEP does not capture any fires between 0 to 5 acres, while almost 20% of the burns in permit records are in this size range.

The comparison of state totals indicates that, in Georgia, both BBEP and GFED4s underestimate the burned areas compared to the permit records (Figure 2-2). GFED4s only accounts for 7.6% and 11% of the burned areas in permit records while BBEP accounts for 15% and 44% in the first four months of 2015 and 2016, respectively. All datasets show an increase in burned areas in Georgia from 2015 to 2016. In Florida, BBEP and GFED4s also underestimate the burned areas with respect to the permit records. GFED4s' state total burned areas are 19% and 16% of those in permit records in 2015 and 2016, respectively. BBEP's are 15% and 75% of burned areas in permit records. Permits and GFED4s show a decrease in burned area from 2015 to 2016 while BBEP shows a large increase.

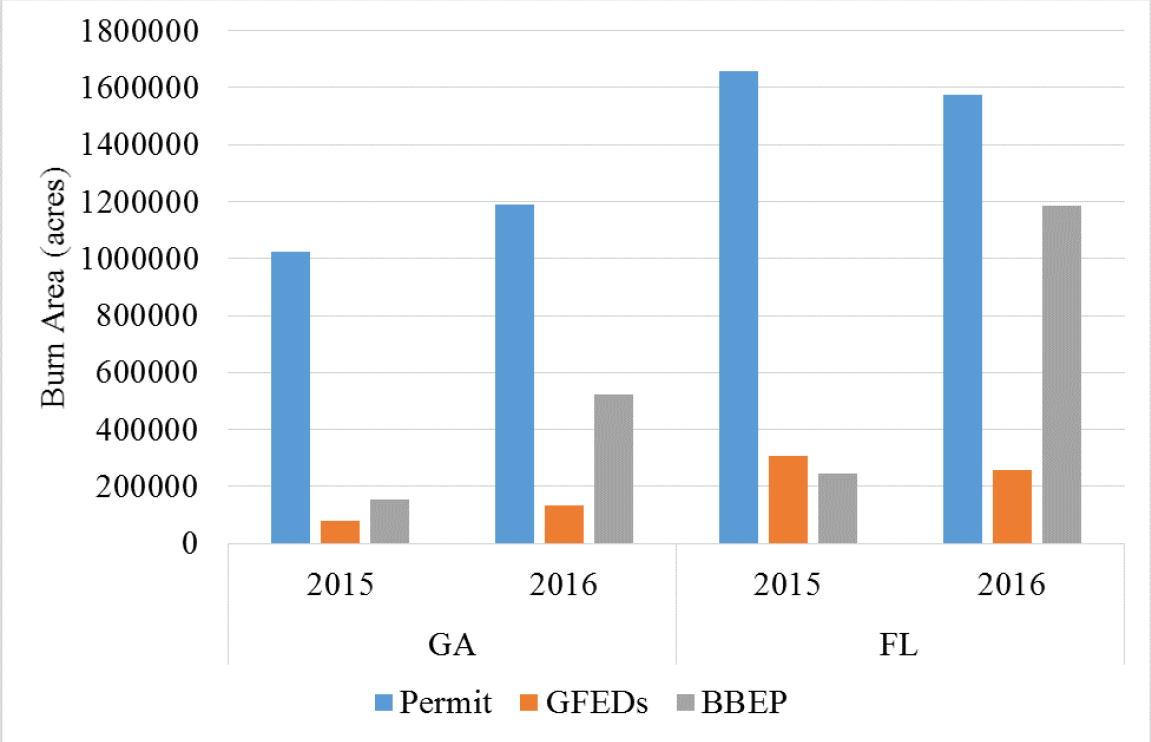

Figure 2-2 Comparison of state total burned areas from permit records, Biomass Burning Emission Product (BBEP), and Global Fire Emissions Database (GFED4s) in Florida and Georgia for the first 4 months of 2015 and 2016.

Total burned areas in permit records of Georgia from January to April in 2015 and 2016 show that most of the burn activity takes place in the southwest portion of the state

(Figure A - 1). The dominant burn types here are land clearing and hazard reduction burns. Neither GFED4s nor BBEP captures the level of burned area in permit records (the scale is 10 times larger for 2015 and 2016 permits in Figure A - 1). BBEP is in better agreement than GFED4s with the permits in terms of the burn locations. In Florida, the largest burned areas in permit records are in the panhandle and the southcentral counties (Figure A - 2 where the scale is five times larger for 2015 permits and four times larger for 2016 permits). The dominant burn type in Florida is agricultural burn. Compared to 2015 burned areas in permit records, most counties have a decrease in total burned area in the first four months of 2016. However, BBEP shows an increase of burned area in 2016 compared to 2015. Both GFED4s and BBEP underestimate the total burned areas for most counties. We also compared the fires between Hazard Mapping System Fire and Smoke Product (HMS, a product which shows the detected hot spots and smoke plumes indicating possible fire locations by combining human analysis with the satellite data) with BBEP and found some differences but nothing that might explain the large difference from the permit data. Prescribed burns are typically ignited around noon and put out before sundown. The flaming phases are typically very short (1-2 hours). If the sky is overcast by clouds during these periods, the satellites cannot detect the fires. Finally, we extracted the 2015 Georgia burned area data from BAECV and compared with the permit records for the whole year (Figure A - 3). The results are similar to the other two satellite products: while BAECV captures the spatial patterns, it underestimates the total burned areas.

The potential uncertainty due to geolocation is reduced in our county-by-county analysis. While the uncertainty of geolocations is a concern (more so in Georgia where some non-standard addresses could not be converted accurately to geolocations than in

Florida where the coordinates were available for all the burns), the burn locations should be within the same county due to permitting requirements and practices. The discrepancies between the satellite and the permits in the county-wise analysis are most likely due to the inability of the satellites to detect the burns. Prescribed fires in Georgia and Florida are usually both low intensity and under the tree canopy making satellite detection difficult. In addition, the frequent presence of clouds in the region during the active burn season further obscures the satellites' view (Gibson and Vonder Haar 1990; Connell et al. 2001).

In addition, we compared the district total burned areas for 11 fire districts in Georgia (Figure A - 4a) and 15 fire districts in Florida (Figure A - 4b). For the first 4 months of 2015, BBEP is more correlated with permit record burned areas than GFED4s (Figure A - 5 and Figure A - 6). There is a strong correlation between BBEP and GFED4s burned areas because both of them use data from MODIS. The correlation is larger for district totals compared to the county totals in Georgia but the opposite is true in Florida.

2.3.3 Comparison of BBEP daily total and GFED4s monthly total burned areas with permit data

The permit record and BBEP daily state total burned areas are correlated in the first four months of 2015 in Georgia and Florida with r^2 equal to 0.57 and 0.66, respectively (Figure 2-3). However, they are not correlated as strongly in 2016 (r^2 = 0.29 and 0.14), because of the days when BBEP state total burned areas are larger than those from permit records. The slope of all regression lines are smaller than 1.0 and in 2015, the year with good correlations between permit records and BBEP, the slopes are smallest with values equal to 0.17 and 0.15 for Georgia and Florida, respectively. These slope values and the

scatter plots of Figure 2-3 show that BBEP may underestimate the burned area by more than 80%. Note that permit records only account for prescribed burns while satellites cannot differentiate prescribed burns from wildfires. However, our investigations showed no major wildfire incidences on those days.

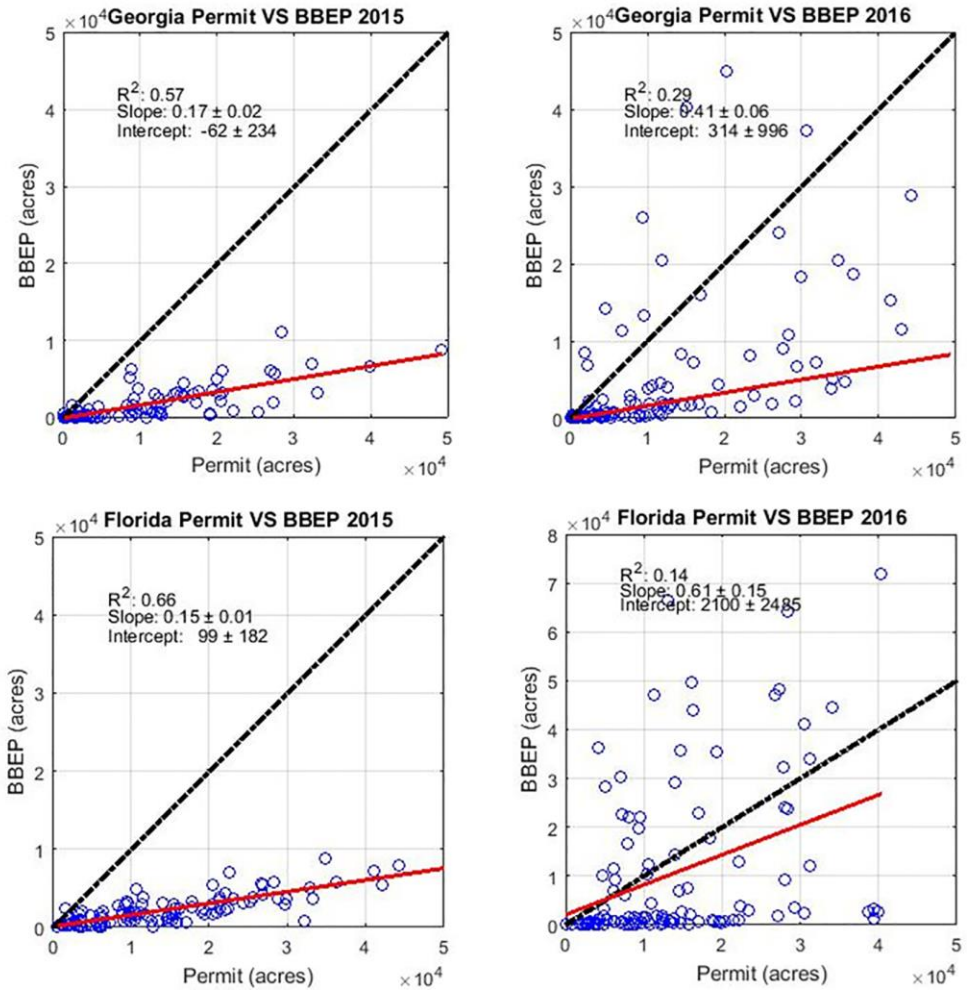

Figure 2-3 Comparison of daily state total burned areas for the first 4 months of 2015 and 2016 in Georgia (top row) and Florida (bottom row): Biomass Burning Emission Product (BBEP) versus permit record data.

The correlation of log-transformations is strong ($R^2 \geqslant 0.63$) in both years for a near-linear relationship between the permit and BBEP burned areas as implied by a slope close to unity (0.74 in 2015 and 0.88 in 2016, Figure A - 7). The similarity in the slopes

and intercepts suggests a similar population for the two years. Therefore, the 2016 results are not really that different, though they are visually different because of the few times BBEP have values above the 1:1 line. When we plotted histograms of the residuals of both the untransformed and transformed data sets for comparison, we found that the log-transformation residuals are more normally distributed than the untransformed residuals (Figure A - 8). In addition, we calculated the Spearman's R between permit records and BBEP for Georgia in 2015 and 2016. The Spearman's R is 0.81 and 0.84 for 2015 and 2016, respectively, while the p-values are much less than 0.05.

Considering GFED4s only has the monthly total burned areas, there are only four points per year in the comparisons between permit records and GFED4s (Figure 2-4). There is strong correlation between GFED4s and permit record burned areas both in Georgia and Florida for 2015 (r^2 = 0.61 and 0.42) as well as 2016 (r^2 = 0.53 and 0.89).

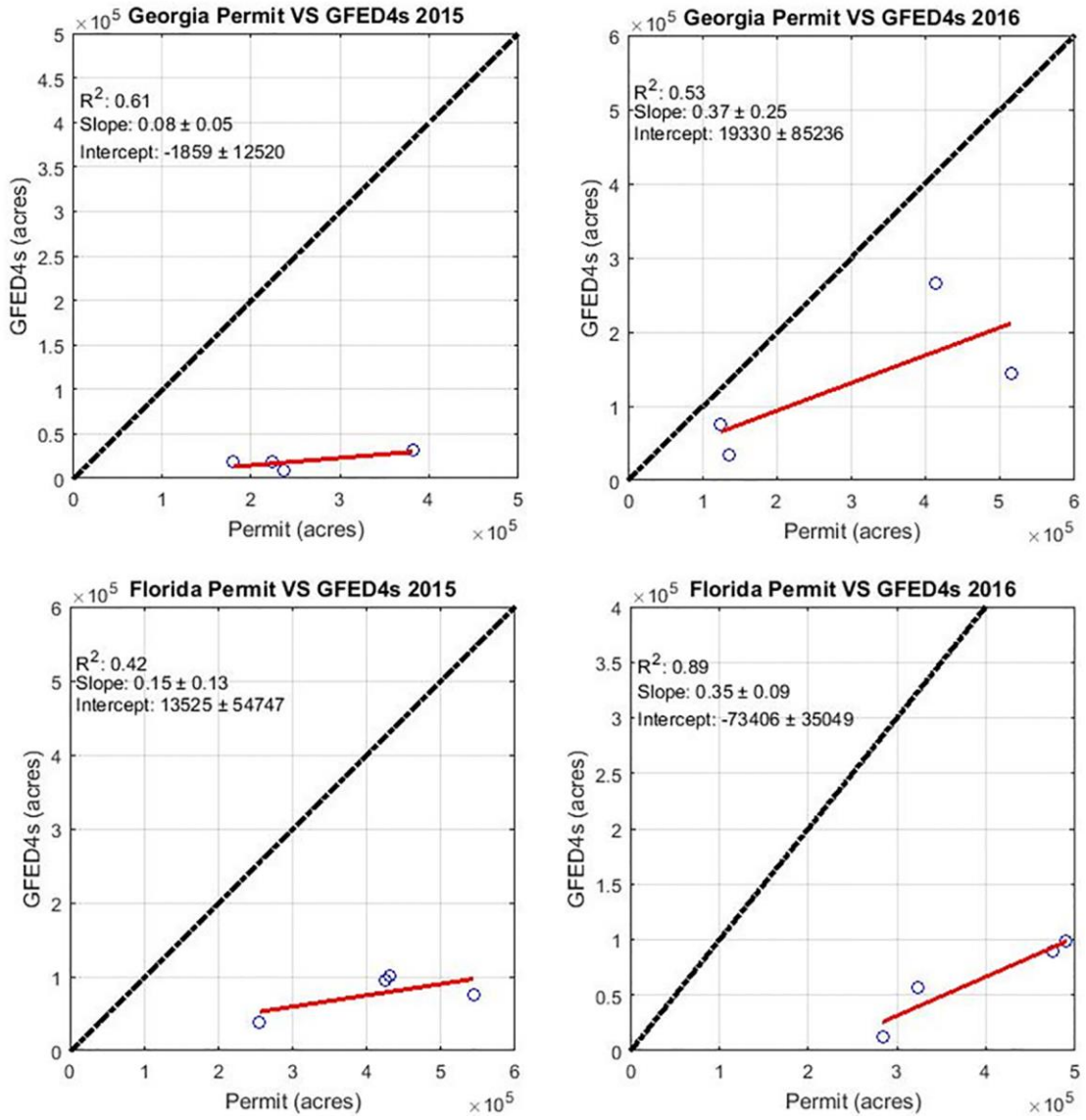

Figure 2-4 Comparison of monthly state total burned areas for the first four months of 2015 and 2016 in Georgia (top row) and Florida (bottom row): Global Fire Emissions Database (GFED4s) versus permit record data.

2.3.4 Comparison of daily fire counts between permit record data and BBEP

BBEP has difficulty detecting the prescribed burns according to the comparison of BBEP fire counts with the burn counts in permit records (Figure A - 9). In Georgia, although the correlations between the daily counts are good (r = 0.71 for 2015 and 0.73 for 2016), BBEP fire counts are only about one-tenth (2015) and one-twelfth (2016) of permit

record burn counts. In Florida, the average daily number of burns from permit records, 120, is four times larger than the average daily number of fires from BBEP in 2015. In 2016, the average daily count from permit records, 132, is more than four times larger than that from BBEP.

2.3.5 Special days analysis in Florida

Because Florida permit record data has latitude and longitude information for all the burns we were able to perform spatial comparisons between permit record burns and BBEP fires. This analysis focused on special days. Uncertainties in this spatial analysis include possible human errors in converting the permit address to latitude and longitude and the limited resolution of the satellites.

For 17 March 2016, although state total BBEP and permit record burned areas are almost the same in Florida (~1.40×10^4 acres), the locations of the fires are quite different (Figure 2-5). In particular, BBEP fires do not agree with the small burns permitted in South Florida; there are only a few small fires detected by the satellite and their locations are different from those of the permitted burns. In addition, larger fires seen by the satellite in South Central Florida are in different locations, even different counties.

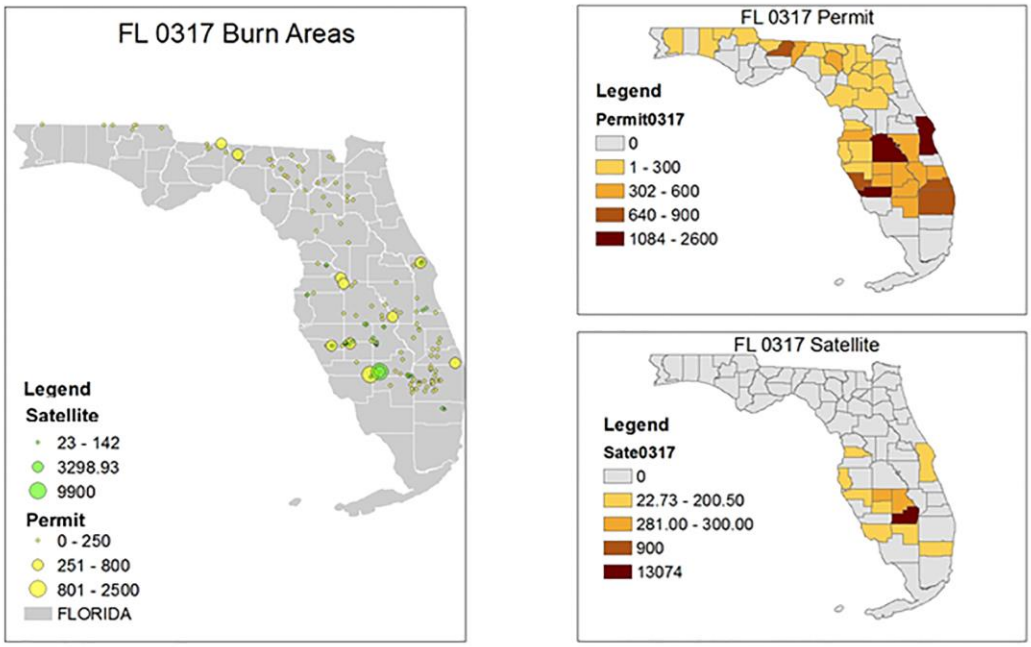

Figure 2-5 Spatial comparison between permitted burns and Biomass Burning Emission Product (BBEP)-detected fires in Florida on 17 March 2016: Individual burned areas (ha) (left) and county-total burned areas (ha) (right).

For 18 March 2016, there is an overlap of permitted burns and BBEP-detected fire locations; however, BBEP detects one very large fire (~10,000 acres) in South Central Florida while the permit database implies many small fires in that area (Figure 2-6). The county total burned area for the satellite-detected fires is much larger than the one for permitted burns in Charlotte County, Florida. BBEP may be interpreting several small fires as one big fire, as previously reported in the literature for other satellite products (Kukavskaya et al. 2013). This would lead to an overestimation by the satellites of the burned area for small fires such as prescribed burns.

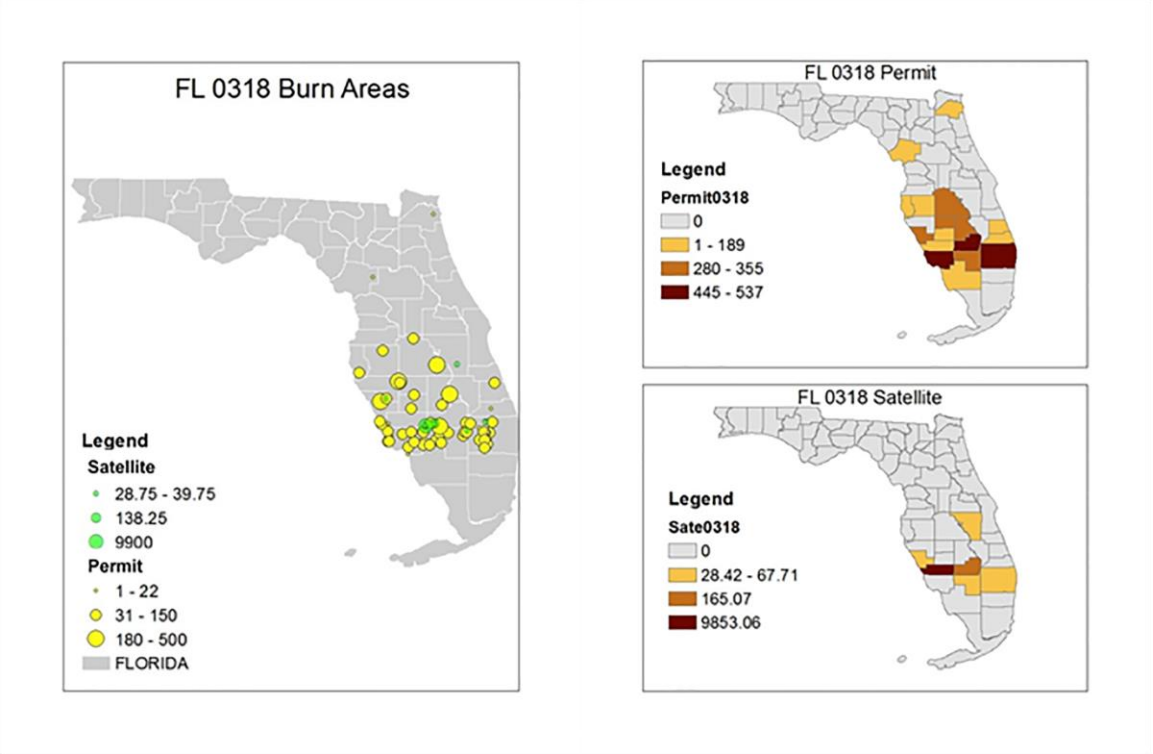

Figure 2-6 Spatial comparison between permitted burns and BBEP-detected fires in Florida on 18 March 2016: Individual burned areas (ha) (left) and county-total burned areas (ha) (right).

2.3.6 Sugarcane burn comparison in Florida

The dominant burn type in Florida's Glades, Hendry and Palm Beach Counties is agricultural sugarcane burn (Figure A - 10). Considering the high frequency of sugarcane burns and that they are conducted in open fields, satellites should be efficient in detecting this type of burns. Taking Palm Beach County as an example, there is no correlation between the burned areas of permitted sugarcane burns and BBEP-detected fires in either 2015 or 2016 (Figure A - 11). However, the red dots in the comparisons of district-wise and county-wise total burned areas in Figure A - 6 that represent District 18, which includes Palm Beach County, and Palm Beach County, respectively, show better than average agreement between the satellite products and permit records. This may be an indication that satellites can detect the sugarcane burns better than other types of burn.

The improvement in sugarcane burn detection is not as obvious according to the comparisons of BBEP-fires with non-sugarcane burns (i.e., all burn types except sugarcane) in Florida (Figure A - 12). The r^2 of the 2015 comparison of daily state total burned areas between BBEP-detected fires in Florida and the permitted non-sugarcane burns is 0.69, slightly larger than that of the comparison with all permitted burns (0.66, Figure 2-3).

2.4 Conclusions

Current satellite-derived products have limitations in estimating the burned areas of small prescribed fires. Comparisons between permit record datasets and two satellite-derived datasets show that satellite products underestimate seriously the burned areas. The BBEP burned areas for 2016 are different from those for 2015, with no correlation between BBEP and permit record data in daily totals; on the other hand, GFED4s has good correlations with permit record data in monthly totals for both 2015 and 2016. Given the limited resolution of the satellite products, from coarser resolution district level to finer resolution county level, the correlation of the satellite burned areas with those of the permit records gets worse.

Satellite-derived products can capture a cluster of fires better than isolated fires, but may misinterpret those small fires together as one big fire, as shown in our special day analyses in Florida. Sugarcane burn is the dominant burn type in three counties in southern Florida. Considering the openness and wide area of the sugarcane plantations, the high frequency of the burns, and the amount of heat produced, sugarcane burns should be detected more efficiently by the satellites. However, burned area comparisons of BBEP

with permit records show only a slight improvement compared to other types of burn. The lack of convergence in sugarcane burns points to systemic problems such as relatively small sizes of those fields with respect to the resolution of geostationary satellites and short durations of the burns that are not in tune with the low frequency of overpasses for Earth-orbiting satellites.

Considering the deficiencies of satellite-derived products, and also the missing data caused by cloudy days, we trust the permit records more. We showed that satellite products vastly underestimated prescribed fire burned areas in the Southeastern U.S. due to their small sizes. Therefore, emission inventories that use satellite-derived burned areas as input should be adjusted accordingly. On the other hand, satellite-derived products need to improve their accuracy in detecting small prescribed fires by taking advantage of new developments. For example, while the current GOES pixel resolution may be too coarse to resolve small fires, the newly launched GOES-16 Advanced Baseline Imager has 2-km spatial and 5-minute temporal resolution; therefore, it should enable considerably better quality products for detecting and monitoring prescribed burning fires.

2.5 Acknowledgment

This publication was made possible in part by funding from NASA Applied Sciences Program under grant numbers NNX11AI55G and NNX16AQ29G, U.S. EPA Science to Achieve Results Program under grant number RD83521701 and the Joint Fire Science Program under grant number 16-1-08-1. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the supporting agencies. Further, The U.S. Government does not endorse the purchase of any commercial products

or services mentioned in the publication. The prescribed fire permit records of Georgia and Florida are available from the Georgia Forestry Commission and Florida Forest Service separately upon request. The BBEP data were obtained from http://www.ospo.noaa.gov/Products/land/bbep2/. The GFED4s data are available from http://www.globalfiredata.org/data.html. We thank:

- Mr. Mark Ruminski of NOAA/NESDIS Satellite Analysis Branch for HMS data
- Drs. Yang Chen and James Randerson of UC Irvine for GFED4s data
- Mr. John K. Fish of Florida Forest Service and his staff for Florida open burn authorization data.
- Mr. Daniel Chan and Mr. Frank Sorrels of Georgia Forestry Commission for Georgia prescribed burn permit data
- Dr. Xiaoyang Zhang of South Dabota State University for BBEP data
- Dr. Shobha Kondragunta of NOAA/NESDIS for BBEP data

CHAPTER 3. APPORTIONING EMISSION SOURCE-GROUP IMPACTS AMONG INDIVIDUAL SOURCES THROUGH DISPERSION MODELING: APPLICATION TO PRESCRIBED FIRES

Abstract

As a preferred land management tool to decrease the particulate matter (PM) emissions and damages from wildfires, there is an increasing trend to have more prescribed burning activity in the U.S. However, prescribed burning is also a prominent source of PM. A novel source apportionment method (Dispersive Apportionment of Source Impacts (DASI)) has been developed and applied to split the combined prescribed fire impact from the chemical transport model (CTM) by using fields from a dispersion model. The results show that DASI works well with large and small emission fires that do not have too much interaction with other fires when comparing the apportioned fire impacts with single burn impacts simulated by CTM directly. Individual burn impacts obtained by splitting the combined burn impacts from CTM could help local land and air quality managers to decide which burns should be allowed or restricted based on their impacts on public health and air quality in areas of concern. DASI could also be applied to conduct source apportionment by splitting the pollutant concentrations from different sources.

3.1 Introduction

Prescribed burning is a land management tool commonly utilized in the United States (U.S.) to maintain healthy ecosystems and to reduce the risk of catastrophic

wildfires. The southeastern U.S. is the most active prescribed burning area. Florida, Georgia and Alabama lead the nation in acreage burned with around 0.8 million ha, 0.5 million ha and 0.4 million ha respectively in 2017 (In Statista - The Statistics Portal. 2018). Research shows that climate change increases the potential for very large fires in the U.S. (Barbero et al. 2015). While prescribed burning is a useful tool to control wildfires, there is a conflict between increased demand for burning and the desire for better air quality for public health concerns. According to the 2014 U.S. National Emission Inventory (NEI) (US EPA 2014), 14% of $PM_{2.5}$ emissions in the U.S. come from prescribed burning. In the southeastern U.S., around 30% of $PM_{2.5}$ emissions are from prescribed burning while only 3% are from wildfires. A clear understanding of the prescribed burning impact on air quality is important.

Most previous research related to prescribed burning impacts on air quality focus on historic fires. Achtemeier et al. (2012) used Daysmoke, an empirical-statistical plume rise model, to simulate a prescribed fire in the southeastern U.S. to show the feasibility of using Daysmoke to model prescribed burning plume rise. Davis et al. (2015) examined uncertainties associated with estimating fire emissions and their effect on smoke concentrations downwind using Daysmoke. Choi et al. (2007) employed the CALPUFF/CALMET/MM5 modeling system to simulate PM_{10} dispersion from agricultural fires to investigate local and regional air quality impacts. Garcia-Menendez et al. (2014) showed, by conducting sensitivity analyses, that successfully modeling the impacts of fires on air quality with a regional-scale chemical transport model (CTM) depends on correctly allocating the fire emissions in space and time. Considering that prescribed burning requires the issuance of a permit, it can be managed to minimize the

impacts on air quality with the maximum amount of burned acres. Balachandran et al. (2017) investigated the sensitivity of ambient $PM_{2.5}$ to various fire weather forecast variables. The results show that prescribed burning decisions should be based on the forecasts released on the morning of the potential burn. Odman et al. (2018) generated a burn activity forecasting decision tree model to estimate prescribed burning emissions and applied it in an air quality forecasting system, which also forecasts fire impacts on air quality using the Community Multiscale Air Quality (CMAQ) (Byun and Schere 2006a) model and Decoupled Direct Method (DDM) (Napelenok et al. 2006), a sensitivity analysis technique for computing sensitivity coefficients simultaneously while air pollutant concentrations are computed. This air quality forecasting system (HiRes2) (Hu et al. 2015; Odman et al. 2018) currently serves most areas in the southeastern U.S. to forecast primary and secondary air pollutant concentrations ($PM_{2.5}$ and O_3) one day in advance. The system could help land managers and allow air quality managers to consider the impacts of burns when predicting the Air Quality Index (AQI) for the next day. Permits may be restricted when an exceedance is imminent and applications that cannot be accommodated on that day may be rescheduled for a future date to burn under more favorable meteorological conditions.

The output prescribed fire impact from the HiRes2 forecasting system is the combined impact of all the fires in the domain. When there are too many fires close to each other, it is difficult to distinguish which fires have a larger impact; therefore, it is not always possible to determine where to issue fewer permits to avoid air quality issues. Considering there are hundreds of burns in a state like Florida or Georgia every day during the burning season, computing the impact of every single burn with CMAQ-DDM would require

massive computational resources. A more practical approach is needed for use by the state agencies in their daily prescribed burn permitting operations.

In this paper, we will describe an efficient method to split the total prescribed fire impact from CMAQ-DDM into individual burn impacts by incorporating dispersion modeling. The method will be applied to the simulation of forecast prescribed burn impacts in South Georgia. The results will be evaluated by comparing the apportioned impacts with single burn impacts simulated by CMAQ-DDM directly. Individual burn impacts obtained by apportioning the combined burn impacts from CMAQ-DDM could help local land and air quality managers to decide which burns should be allowed based on the magnitudes of adverse effects on public health or air quality in areas of concern. The apportionment method is general and can be applied to emission sources other than the fires.

3.2 Method

We developed a new method to split the impact of a source group into its constituents: Dispersive Apportionment of Source Impacts (DASI). The source group may be an entire emission sector (e.g., EGU sector), an ensemble of sources with specific characteristics (e.g., coal-fired power plants (CFPPs)) or simply a cluster of emission sources in a geographic area (e.g., CFPPs in Georgia). In each case, the source group consists of individual sources with a specific address, unique properties, and different emissions. Chemical transport modeling with Eulerian grid models starts with the processing of emissions from these individual sources. During the gridding process, individual source emissions are mixed with emissions from other sources in the same grid

cell. The air pollutant concentration fields produced by the chemical transport models (CTMs) are the combined result of all the emissions sources.

When CTMs are used for the design of emissions control strategy design, the desired output is the contribution of individual sources or groups of sources to air pollution levels. Quantitative information on the contributions of the sources can help determine the level of control required to achieve the desired air quality. The use of Eulerian grid CTMs for the analysis of single source impacts is not very common (Bergin et al. 2008); typically, they are used for the analysis of the impacts from a source group. Eulerian grid CTMs such as CMAQ and CAMx (Environ 2016) are equipped with several tools for these kinds of analyses. For example, DDM calculates the sensitivities of pollutant concentrations to changes in emissions from user-specified sources as derivatives of concentration fields with respect to emissions (Napelenok et al. 2006). These sensitivities can be converted to contributions using Taylor series expansion (Hakami et al. 2003). Source apportionment tools such as OSAT (Dunker* et al. 2002) and PSAT (Wagstrom et al. 2008) calculate the contributions more directly but by compromising some accuracy. In the absence of such methods in a Eulerian grid CTM, one can always employ the brute-force method by conducting a second simulation with reduced or zeroed-out emissions from a source group. The difference between the base case results and reduced (or zeroed-out) emission simulation can be attributed to that source group as its contribution to the air pollution levels (Odman et al. 2019).

The time required for the computation of source contributions with the above methods and the associated cost can be very large, especially for the brute-force method. The brute-force method is the most time consuming and most expensive method. Although

DDM and source apportionment was developed to shorten the computation time and lower the cost, they can be restrictive for analyzing a large number of individual sources. It is for this reason that source impacts are almost never analyzed at the individual source level for all the sources. DASI is designed to work for a large number of individual sources. The idea behind DASI is to split the impact of a source group computed with Eulerian grid chemical transport modeling into the impacts of individual sources via dispersion modeling. The computations of dispersion models such as Gaussian plume and Lagrangian puff or particle models are far less involved than those of Eulerian grid CTMs since they do not include detailed atmospheric chemistry. Therefore, it is feasible to execute these dispersion models for a large number of individual sources in a short amount of time at a very low cost. The pollutant concentration fields predicted by dispersion models downwind of each individual source can then be used together to apportion the total impact of the source group as predicted by the Eulerian CTM.

3.2.1 DASI Equations

3.2.1.1 Splitting source-group impact from Eulerian CTM using dispersion modeling

The individual impact of source p is obtained from the total impact of the source-group as follows:

$$s_{i,j}^{p} = \frac{m_{i,j}^{p}}{\sum_{p=1}^{N} m_{i,j}^{p}} \times S_{i,j}. \qquad (1)$$

Here,

s^{p} is the individual impact of source p as vertical column mass of pollutant (in µg),

m^p is the vertical column mass of pollutant from dispersion modeling of source *p* (in µg),

S is the source-group impact from Eulerian CTM as a vertical column mass of pollutant (in µg),

i and *j* are the horizontal column and row indices for the vertical column, and

N is the number of individual sources in the group.

The reason for using vertical column totals is that there is usually a large difference between the vertical distributions of pollutants in Eulerian CTM and Gaussian/Lagrangian dispersion models as will be shown later. The vertical column totals, as the mass of pollutant, are calculated from grid concentrations as follows:

$$S_{i,j} = \sum_{k=1}^{K} C_{i,j,k} \times A_{i,j} \times \Delta H_k \qquad (2)$$

$$m_{i,j}^{p} = \sum_{k=1}^{L} q_{i,j,k}^{p} \times A_{i,j} \times \Delta h_k \qquad (3)$$

Here,

C is the source-group impact (as pollutant concentration) from Eulerian CTM (in µg/m^3),

q^p is the pollutant concentration from dispersion modeling of individual source *p* (in µg/m^3),

A is the grid cell area in Eulerian CTM (in m^2),

ΔH is the layer height in Eulerian CTM (in m),

Δh is the layer height in Gaussian/Lagrangian dispersion model (in m),

k is the layer index, and

K and L are the number of vertical layers in Eulerian CTM and Gaussian/Lagrangian dispersion model, respectively.

Note that the vertical column pollutant mass from the dispersion model is summed over the Eulerian CTM grid-cell area in Equation 3. Typically, the horizontal grid used in the dispersion model is different and of finer resolution than the one in the Eulerian CTM. This necessitates mapping of the concentration field from the dispersion model's horizontal grid to the Eulerian CTM's grid. We find all the dispersion grids which are finer (e.g. 1km, 500m) that located in one Eulerian CTM grid (e.g. 4km) and sum up the mass of those dispersion grids to get the mass of that one Eulerian CTM gird. In general, the vertical layer structures of the dispersion model and Eulerian CTM are also different. Note that once the dispersion model's concentration field is mapped onto the Eulerian CTM's horizontal grid, all that needed is to sum up the pollutant mass (or impact as mass) layer-by-layer for each model's vertical columns. This is why the upper limits of summations, K and L, are different in Equations 2 and 3. It is not necessary to go all the way to the top layers of the models, as long as the upper limit layers extend to heights greater than plume heights in each model, which may be different from each other.

3.2.1.2 Matching the horizontal plume extents

Just like the possibility of the vertical extents of the plumes for the same individual source being different in the dispersion model and the Eulerian CTM, the horizontal extents can also be different. In that case $\sum_{p=1}^{N} m_{i,j}^{p}$ in Equation 1 can be zero for certain *i*, *j* for which $S_{i,j}$ is non-zero. Since this would lead to a division by zero, to make Equation 1 usable, it is necessary to match the horizontal extent of the plume in the dispersion model with the one in the Eulerian CTM. This can be achieved by applying artificial diffusion to the column mass fields obtained from the dispersion model. The following equation is applied iteratively to all $m_{i,j}^{p}$ until no $\sum_{p=1}^{N} m_{i,j}^{p}$ is zero for non-zero $S_{i,j}$:

$$m_{i,j}^{p\ *} = m_{i,j}^{p} + D\left(m_{i-1,j}^{p} + m_{i,j-1}^{p} - 4m_{i,j}^{p} + m_{i,j+1}^{p} + m_{i+1,j}^{p}\right). \qquad (4)$$

Here,

D is non-dimensional artificial diffusion coefficient.

We recommend a value of 0.1 for *D*, which achieved the non-zero condition in less than six iterations in the test case described below. The $m_{i,j}^{*}$ obtained from the last iteration can be safely used in Equation 1 to split the source group impact $S_{i,j}$.

3.2.1.3 Extracting ground layer concentration from split total vertical column mass of single fire

The column total impact of source *p*, s^{p}, is distributed to vertical layers *k* by assuming that it has the same vertical profile as the source-group impact, *S*, at column *i*, *j*. First, a mass ratio, *R*, for layer *k* is defined as

$$R_{i,j,k} = \frac{C_{i,j,k} \times A_{i,j} \times \Delta H_k}{S_{i,j}}. \quad (5)$$

Then, this ratio is applied to s^p to get the impact of individual source *p* for layer *k* (as pollutant concentration):

$$c^p_{i,j,k} = \frac{R_{i,j,k}}{A_{i,j} \times \Delta H_k} \times s^p_{i,j}. \quad (6)$$

Here,

c^p is the impact (as pollutant concentration) of individual source *p* (in μg/m^3).

For the surface layer (*k*=1), the impact of individual source *p* is therefore

$$c^p_{i,j,1} = \frac{R_{i,j,1}}{A_{i,j} \times \Delta H_1} \times s^p_{i,j}. \quad (7)$$

3.2.2 Models and modeling domain

Here, the Eulerian CTM is CMAQ, and the dispersion model is the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Draxler and Hess 1997; Draxler and Hess 1998). The emission source group is prescribed fires and each one of the fires is an individual source. The source-group impacts are calculated using DDM, which is embedded in CMAQ. The prescribed fire emissions are calculated using the BlueSky Framework. The modeling domain is the area circumscribed by the black square in Figure 3-1. We focus on four prescribed fires in South Central Georgia near the Florida border on April 27, 2016. One fire is relatively isolated (ID01) while the other three are clustered (ID02~ID04) and their plumes are likely to interact under south-southwesterly winds on

that day. The burned area in each fire is equal (200 acres or 80 hectares), however, the differences in fuel loads lead to different heat fluxes and emissions.

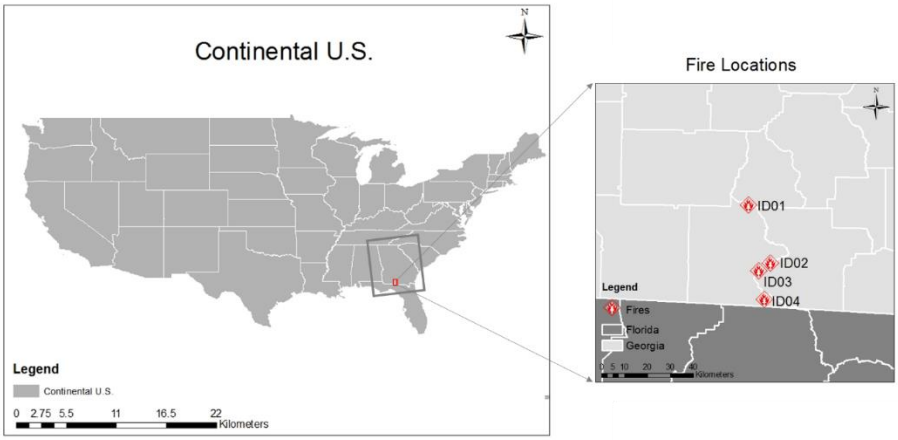

Figure 3-1 Modeling domain (black square) and the locations of the four fires (red flame)

3.2.2.1 Modeling prescribed fire emissions: the BlueSky Framework

The BlueSky Framework (Larkin et al. 2009) is a smoke modeling framework that links a variety of state-of-the-art models of meteorology, fuels, consumption, emissions, and air quality to enable simulations of the cumulative smoke impacts from fires. The default fuel map is from the Fuel Characteristic Classification System (FCCS) in BlueSky. CONSUME model Version 3 has been used to calculate consumption as the default pathway in BlueSky. Emissions Production Model (EPM), Fire Emissions Production Simulator (FEPS) and First Order Fire Effects Model (FOFEM) are integrated into the BlueSky framework to calculate emissions. We assume the fire starts at 11 am (EST) and the flaming and smoldering phases continue for six hours.

In the process of applying DASI, one must pay attention to how emissions are input into the models. It is very important that the same mass of pollutant emitted in the Eulerian CTM is injected into the Gaussian/Lagrangian dispersion model. CMAQ emission inputs are instantaneous flux values and the model interpolates between two flux values in the input file to calculate the emitted pollutant mass for each time step. In contrast, HYSPLIT inputs are time-averaged emission fluxes and the model uses the same flux while advancing from the time stamped on one flux value to the time on the next one. This difference must be kept in mind when preparing the emission inputs for the same models. For example, to inject 1 kg of pollutant starting from 10:00 a.m. until 11:00 a.m., the emissions input file for CMAQ must have two records: 0 kg/hr for 10:00 a.m. and 2 kg/hr for 11:00 a.m. This will inevitably result in an additional emission of 1 kg from 11:00 a.m. to 12:00 p.m. even if the emission record for 12:00 p.m. is 0 kg/hr. To match this, the following records must be in HYSPLIT's emissions input file: 1 kg/hr for 10:00 a.m., 1 kg/hr for 11:00 a.m. and 0 kg/hr for 12:00 p.m.

3.2.2.2 Meteorology and chemical transport modeling: HiRes2 Air Quality and Source Impacts Forecasting System

HiRes2 is a regional forecasting system that has been in operation since November 2014 to provide local air quality and source-impact forecasts for the southeast U.S. The system uses the Weather Research and Forecasting model (WRF, version 3.6) for forecasting meteorology, the Sparse Matrix Operator Kernel Emissions model (SMOKE) for gridded emissions and the Community Multiscale Air Quality model (CMAQ, version 5.0.2) for chemistry and transport. All non-fire emissions come from 2011 National Emission Inventory. The horizontal resolution of HiRes2 over the modeling domain in

Figure 3-1 is 4km × 4km. There are 13 vertical layers: 0, 20m, 40m, 81m, 162m, 326m, 663m, 1053m, 1554m, 3291m, 5485m, 7989m, 11643m, and 20201m. The source-impacts of various emission source-groups including prescribed fires are calculated by DDM available in CMAQ. We ran CMAQ-DDM with two different emission scenarios. In the first scenario, emissions of all fires (ID01-04) are combined together. In the second scenario, the model was run separately with the emissions from each individual fire.

3.2.2.3 Dispersion modeling with HYSPLIT

Goodrick et al. (2013) reviewed the dispersion models used in modeling smoke transport from wildland fires, including Gaussian plume and Lagrangian puff or particle models. Puff models provide a significant advantage over Gaussian plume models as they can effectively deal with time-varying meteorological conditions and complex terrain, two limitations of plume models. HYSPLIT is a Lagrangian puff model commonly used to estimate smoke concentrations downwind from fires. Rolph et al. (2009) developed a smoke forecasting system using HYSPLIT together with the BlueSky emission algorithm and the Hazard Mapping System (HMS) satellite analysis. The Florida Fire Management Information System (Brenner and Goodrick 2005) also used HYSPLIT to estimate smoke plume movement and ground-level impacts on $PM_{2.5}$ concentrations. We used the HYSPLIT 4 (Windows version) to simulate the primary $PM_{2.5}$ concentration fields downwind from each individual fire. The meteorology inputs are derived from the Weather Research and Forecasting model (WRF, version 3.6), which is also used in the HiRes2 air quality forecast system for generating meteorological inputs to CMAQ. The horizontal grid resolution is 1km × 1km. The vertical height has been separated into nine layers: 0, 20m, 40m, 80m, 160m, 320m, 640m, 1000m, 1500m, and 2000m. The input emissions come

from BlueSky and are the same as the fire emissions input to the CMAQ model. The 1km × 1km resolution output concentration fields are mapped onto the 4km × 4km CMAQ grid during the application of DASI.

3.3 Results and Discussion

3.3.1 BlueSky emissions and plume height

The hourly emissions (Figure 3-2) shows that even though the burned area is the same, the emissions can be different with different fuel loadings and fuel consumptions. The plume height (Figure 3-3) of the fires is comparable to those reported in previous research (Liu 2014; Davis et al. 2015) with maximum heights between 600m and 1200m.

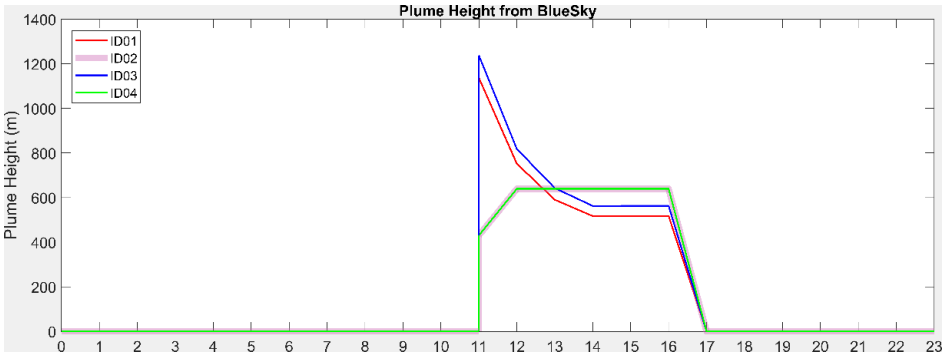

Figure 3-2 Hourly $PM_{2.5}$ emissions from the four burns as estimated by BlueSky

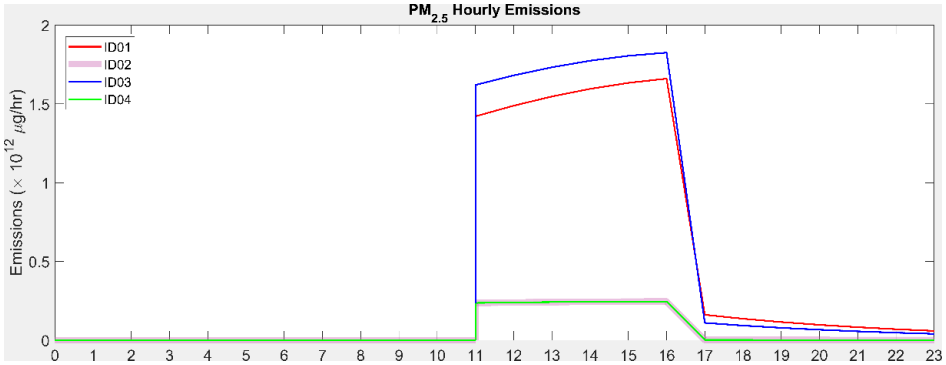

Figure 3-3 Hourly plume heights of the four burns as estimated by BlueSky

3.3.2 Comparison of fire plumes between HYSPLIT and CMAQ-DDM

We performed a test with fire ID01 to find the difference between the $PM_{2.5}$ concentrations predicted by HYSPLIT and the fire impact (as $PM_{2.5}$ concentration) from CMAQ-DDM. In HYSPLIT (Figure 3-4), the highest $PM_{2.5}$ concentrations in the plume are in the upper two layers. However, the CMAQ-DDM result (Figure 3-5) shows that the most concentrated layers are near the ground. The vertically integrated mass (Figure 3-6) shows that the highest mass grid from HYSPLIT (= 3.7×10^{12} μg) is a grid close to the fire location, while from CMAQ-DDM (= 4.5×10^{11} μg), the highest mass is found in the grid where the fire is located. The difference of vertical distribution and the grid of highest mass between HYSPLIT and CMAQ-DDM is mainly because of the difference in the dynamics of those two models and because the layer heights from HYSPLIT and CMAQ-DDM are different. In order to minimize the effect of the difference between the models, the mass of all vertical layers is added together for the following analysis. In the test case (Figure 3-6) we also found that there is less dispersion represented in HYSPLIT compared to CMAQ. We artificially diffused the HYSPLIT plume using Equation 4 to better match the horizontal extent of the fire plume in CMAQ.

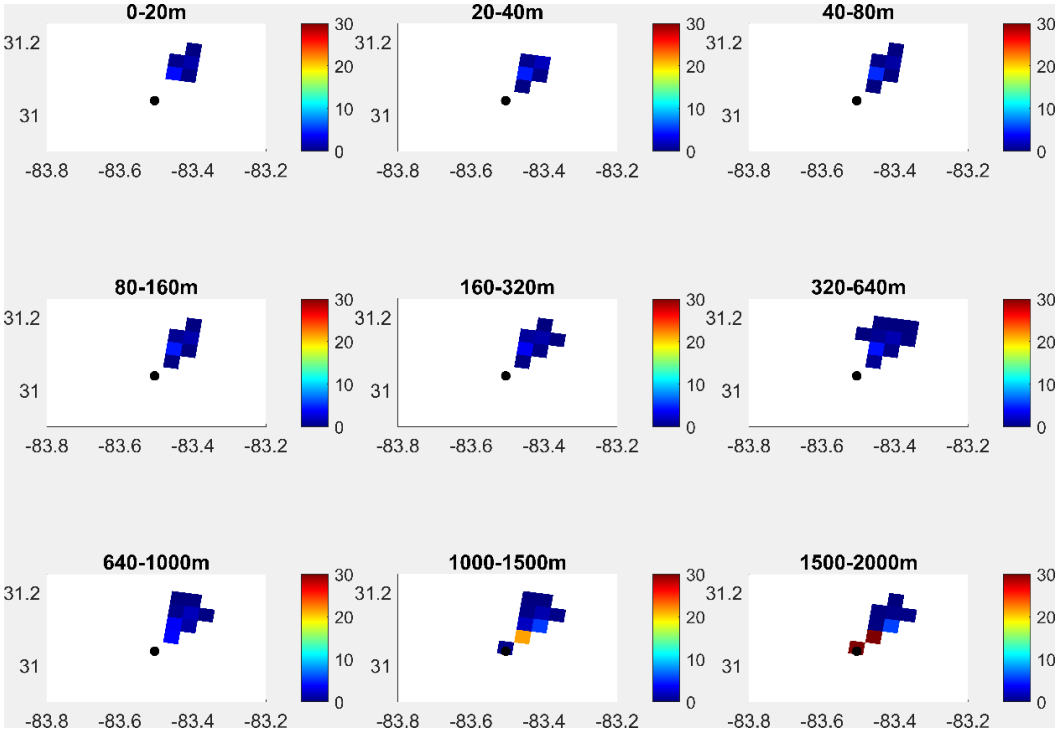

Figure 3-4 Layer concentrations of $PM_{2.5}$ for ID01 from HYSPLIT after the first hour of the burn (mapped onto the 4km × 4km CMAQ grid)

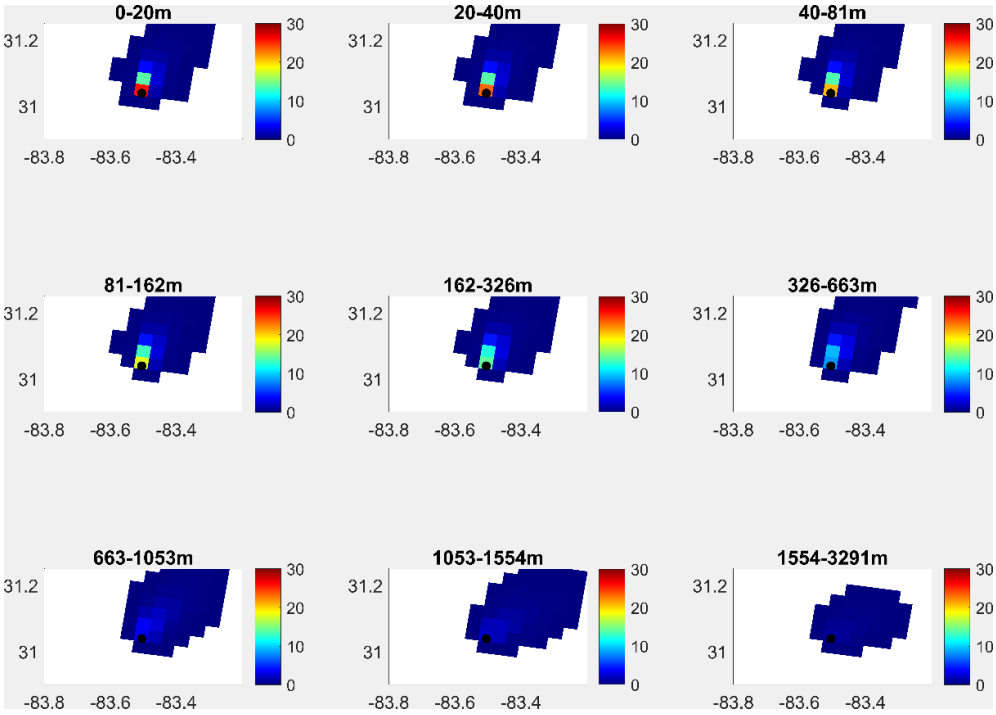

Figure 3-5 Layers concentrations of ID01-related $PM_{2.5}$ from CMAQ-DDM after the first hour of the burn

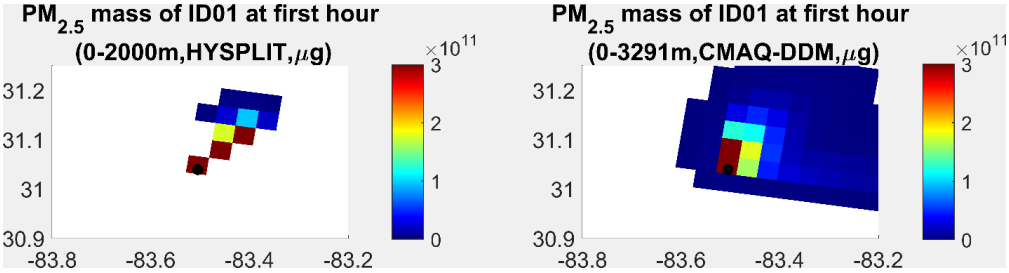

Figure 3-6 Vertical column mass of $PM_{2.5}$ for ID01 from HYSPLIT and CMAQ-DDM after the first hour of the burn

3.3.3 Comparison between combined fire impact and the sum of single fire impacts from CMAQ-DDM

The fire impact ($PM_{2.5}$) based on the combined emissions of all three fires (Figure 3-7a), the sum of the individual impacts of all three fire (Figure 3-7b) and the absolute and relative differences (Figure 3-7c and Figure 3-7d) between those two show that the $PM_{2.5}$ column total concentration from CMAQ-DDM is nearly linearly correlated with fire emissions. The highest concentration grids are 120.56 µg/m^3 (a) and 120.48 µg/m^3 (b) from the two cases. In each case, the highest concentration is from the grid in which ID03 is located. The difference between the concentrations from the two runs is less than 1.3 µg/m^3 (5%). The grids with the largest differences are located near fire ID02 and are impacted by all three fires. This shows there is little non-linear interaction among the three fire plumes.

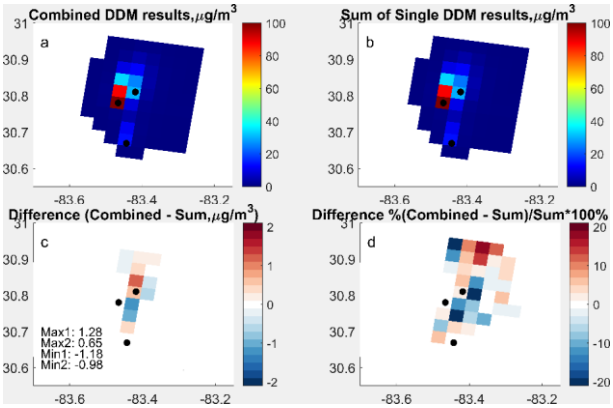

Figure 3-7 Comparison between combined fire impacts and summation of single fire impacts and the absolute and relative differences between those two

3.3.4 Comparison between apportioned individual fire impact and single fire impact from CMAQ-DDM

The process of splitting the combined fire impacts (Equations 1–3) according to the artificially diffused HYSPLIT fields (Equation 4) for fires ID02, 03 and 04 are shown in Figure B - 1, Figure B - 3 and Figure B - 5, respectively. The vertical column mass then converted to obtain the surface layer concentration according to Equations 5-7. For each fire, the highest impacts (as $PM_{2.5}$ concentration) are in the grid cells either the fire is located or the nearby grid cells (Figure B - 2, Figure B - 4 and Figure B - 6; black dot represents the fire to split; purple dots represent the other two fires). When comparing the split fire impact field with the single fire CMAQ-DDM field, the grid with the highest concentration is the same except for ID02. The highest grid-cells concentrations are 0.95 and 0.88 $\mu g/m^3$ from the split fields for ID02, which are 68% and 46% smaller than the concentrations (2.75 and 1.75 $\mu g/m^3$ respectively) from the single fire CMAQ-DDM results. For the other grid cells, the differences are less than 0.5 $\mu g/m^3$. For fire ID03 which has the largest emissions, the split fields and the single fire CMAQ-DDM fields are similar to each other. The highest concentration grid cells are the same for both cases, which are 23.82 (split) and 24.63 (CMAQ-DDM) $\mu g/m^3$. The difference between the split result and single fire CMAQ-DDM result for each grid is less than 2 $\mu g/m^3$. The largest difference (1.94 $\mu g/m^3$) is for the grid in which fire ID02 is located, which may be due to a non-linear interaction among the fires. For fire ID04, which has similar emissions as fire ID02, the differences between the split results and single fire CMAQ-DDM results are small for most grids, especially for the grid in which the fire is located. The difference is only 0.007 $\mu g/m^3$ (<1%). The largest differences are found in the grids in which the other two fires are

located. Even though the three fires have the same area, the difference in emissions results in fire ID03 contributing the most to $PM_{2.5}$ concentration in the combined fields. The split fields match well with the single fire CMAQ-DDM fields. The method is efficient, reliable and could save lots of computational time.

Fire ID02 has less $PM_{2.5}$ emissions and impacts mostly the same grid cells that are impacted by another fire with larger emissions (ID03). This causes a larger difference between the split fire impact and single CMAQ-DDM fire impact for fire ID02 than for the other two fires. The sum of the ground level concentrations from the DASI method is 35% smaller than that from the single fire CMAQ-DDM method (Table 3-1). While for fire ID03 and ID04, total ground level concentrations from the DASI method are 4% and 3% larger than those from the single fire CMAQ-DDM method. For fire ID03, which has the largest emissions, the variance of the different fields for the two methods is quite small, showing that the method works well for large emission fires. The method also works well for fire ID04. Although it has small emissions similar to fire ID02, it doesn't impact many grid cells that are also impacted by the other two fires.

Table 3-1 Statistical information of ground layer concentrations from DASI-split impact (SPLIT), single fire impact from CMAQ-DDM (Single CMAQ-DDM) and their difference (SPLIT – Single CMAQ-DDM) fields

	ID02				ID03				ID04			
	SPLIT	Single CMAQ-DDM	Diff*	Diff * (%)	SPLIT	Single CMAQ-DDM	Diff*	Diff * (%)	SPLIT	Single CMAQ-DDM	Diff*	Diff * (%)
Sum	4.30	6.58	-2.27	-35%	52.06	49.92	2.15	4%	5.94	5.79	0.15	3%
Mean	0.12	0.23	-0.06	-49%	1.13	1.22	0.05	-7%	0.20	0.18	0.005	13%
Variance	0.05	0.34	0.12	-85%	15.60	18.31	0.12	-15%	0.32	0.30	0.01	7%

* Diff: difference between DASI-split single fire impact field and CMAQ-DDM single fire impact field.

Diff (%): Relative difference ((SPLIT-Single)/Single×100%)

3.4 Conclusions

Prescribed burning is becoming an increasing $PM_{2.5}$ source in the southeast U.S. as the potential for wildfire increases. We provided a method to help land and air quality managers identify the most impactful fires under the current weather conditions with ease. The method aims to split the combined impact of a source-group, here prescribed fires, into individual source impacts through dispersive apportionment. The CMAQ-DDM calculated prescribed fire impact from the HiRes2 Air Quality and Source Impacts Forecasting System was efficiently and quickly apportioned to single fire impacts using the HYSPLIT dispersion model. The results show that the method performs well both with large and small emission fires that do not have too much interaction with other fires. The largest inaccuracies with the fire impacts on pollutant concentrations are usually observed in the grids that contain one of the fires and that are also affected by other fires at the same time. Inaccuracies may also be significant at nearby downwind grid cells.

Current source apportionment methods like receptor modeling methods that are based on mass balance analysis do not consider the chemical reactions between species which are important and could bring inevitable uncertainty to the results. Other methods like sensitivity analysis methods using a chemical transport model (CMAQ-DDM) could provide appropriate results but require too many computational resources. Our method (DASI) could not only be applied to split the combined fire impacts but could also be used to split the impacts on pollutant concentrations from different emission sources with the limited cost of time and space.

3.5 Acknowledgment

This publication was made possible in part by funding from U.S. EPA Science to Achieve Results Program under grant number RD83521701, Joint Fire Science Program under grant number 16-1-08-1, and NASA Applied Sciences Program under grant number NNX16AQ29G. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the supporting agencies. Further, The U.S. Government does not endorse the purchase of any commercial products or services mentioned in the publication. We thank:

- Dr. Momei Qin of EPA to provide the code convert 1km HYSPLIT fields to 4km CMAQ-DDM fields.

CHAPTER 4. APPLICATION AND EVALUATION OF A LOW-COST PM SENSOR TO QUANTIFY THE IMPACTS OF PRESCRIBED BURNING ON AIR QUALITY IN SOUTHWESTERN GEORGIA

Abstract

Prescribed burning is a prominent source of $PM_{2.5}$ in the southeastern U.S. and exposure to prescribed fire smoke is a health risk. As the demand for burning increases and stricter controls are applied to other pollution sources, prescribed burning emissions will be responsible for an increasing fraction of $PM_{2.5}$ concentrations. In order to quantify the effect of prescribed burning on air quality, low-cost PM sensors have been used to measure the $PM_{2.5}$ concentrations in southwestern Georgia. Here, the feasibility of using low-cost sensors as a supplemental measurement tool is evaluated by comparing the measured $PM_{2.5}$ concentrations with reference instruments (β-attenuation monitors). A chemical transport model (Community Multiscale Air Quality (CMAQ)) was also used to simulate the contribution of prescribed burning on $PM_{2.5}$ concentrations using the decoupled direct method to understand the impact of prescribed burning on the local air quality and was compared to observations using both the low-cost sensors and reference monitoring. The results show that the severe impact of prescribed burning on local air quality and public health may be missed due to the dearth of regulatory monitoring sites in Southwestern Georgia. Low-cost PM sensors can be used to detect prescribed fire impacts and provide spatial information for integration with and evaluation of air quality models. Further, $PM_{2.5}$ concentrations in Southwestern Georgia are not homogeneous and the spatial variation is

not captured even with a 4-km horizontal resolution in air quality model simulations. In the future, observations from a dense network of low-cost sensors could be fused with the model simulated $PM_{2.5}$ fields to provide accurate estimates of exposures to smoke from prescribed burning.

4.1 Introduction

Particulate matter, a major component of air pollution, is associated with increased incidences of cardiovascular and respiratory disease (Brook et al. 2004). Over 4 million deaths worldwide (8% of total global mortality) were caused by exposure to outdoor $PM_{2.5}$ (particulate matter with an aerodynamic diameter less than $_{2.5}$ μm) in 2015 (Forouzanfar et al. 2016). In the U.S., $PM_{2.5}$ is the environmental risk factor with the largest health burden and 6th largest mortality risk overall (Cohen et al. 2017).

Wildland fires, including wildfire and prescribed burning, are a major source of $PM_{2.5}$. 30% of $PM_{2.5}$ emissions in the U.S. comes from wildland fires according to the 2014 U.S. National Emission Inventory (NEI) (US EPA 2014). Prescribed burning is a land management tool practiced to reduce wildfire risks, control pest insects and disease, and recycle nutrients back to the soil. 14% of $PM_{2.5}$ emissions in the U.S. and 30% in the southeastern U.S. originate from prescribed burning. Georgia is one of the most active prescribed burning states in the Southeastern U.S. with a total burned area around 550,000 ha in 2016 (Huang et al. 2018b). As stricter controls are applied to other pollution sources, prescribed burning emissions will provide an even larger contribution to $PM_{2.5}$ concentrations.

$PM_{2.5}$ increases significantly under the influence of fire smoke (Reisen et al. 2011). Rappold et al. (2017) found that over 40% of Americans live in areas with a moderate or high contribution of wildland fires to ambient $PM_{2.5}$ concentrations based on model simulations. Researchers have also found associations between fire smoke and respiratory morbidity (Dennekamp and Abramson 2011), cardiovascular disease (Haikerwal et al. 2015) and additional premature deaths (Fann et al. 2018). A better understanding of the contributions of prescribed burning to air pollution and its impacts on public health is important, especially to the local populations affected by prescribed burning directly. Finer resolution exposure fields are often generated by fusing observation and model simulation, while model simulation provides spatial information to the fields for health studies (Huang et al. 2018a). However, the sparse distribution of monitoring sites limits the information available to understand the impact of prescribed burning on local air quality and public health at fine scales. Deployment of inexpensive devices to measure ambient pollutant concentrations could provide better resolved spatial information to improve the accuracy of exposure fields to quantify prescribed burning's impacts on air quality and public health.

Air pollution sensors that are lower-cost, portable and easy-to-use have been widely used as a supplemental tool to measure ambient concentration to provide high-resolution data in near real-time (Snyder et al. 2013; Kumar et al. 2015; Jovašević-Stojanović et al. 2015; Rai et al. 2017). Previous research evaluating different types of low-cost sensors in both laboratory and field studies have shown varying performance between the sensor measurements and reference instruments (Gao et al. 2015; Johnson et al. 2016; Kelly et al. 2017; Han et al. 2017; Johnson et al. 2018; Zheng et al. 2018). Gao et al. (2015) tested the performance of a low-cost sensor in high concentration urban environments in Xi'an,

China. The results show that the low-cost sensor could identify the area with the highest average $PM_{2.5}$ concentration in the study. Lower R^2 values were obtained at lower ambient concentrations. Han et al. (2017) evaluated a low-cost sensor and compared its measurement with a GRIMM Mini Laser Aerosol Spectrometer, Model 11R in an urban residential area of Houston, Texas that shows good correlation between those two measurements. They also mentioned that RH significantly changes the association between the low-cost sensor and the official measurement. Relative humidity (RH) and temperature are important in the calibration of measurements of the PM concentration. Zheng et al. (2018) evaluated a low-cost PM sensor in both low concentration suburban regions (Durham and Research Triangle Park, North Carolina) and a high concentration urban location (Kanpur, India). Low-cost sensor performance improved as ambient $PM_{2.5}$ increased. They also pointed out that β-attenuation monitors (BAM) may not be ideal for testing low-cost PM sensors at low concentrations. Although the use of low-cost sensors in the field of wildland fire is limited, those studies also show that low-cost sensors have better performance in high concentration environments; this is advantageous for measuring $PM_{2.5}$ concentrations impacted by fires. Kelleher et al. (2018) developed a low-cost $PM_{2.5}$ sampler and evaluated its performance as a smoke-monitoring tool during a prescribed burning activity in Colorado, from September 8 to 17, 2016. The regression between the low-cost sampler and reference instrument (BAM) found good agreement (R^2=0.92). Gupta et al. (2018) deployed a low-cost air quality monitor network in California to quantify the impact of wildfires during October 2017. They also found that low-cost sensors are useful in developing statistical models to convert aerosol optical depth into $PM_{2.5}$.

In this paper, we will evaluate the feasibility of a low-cost sensor (Plantower PMS 3003) by comparing the measured $PM_{2.5}$ concentration with a reference instrument in Southwestern Georgia. We also use a chemical transport model (Community Multiscale Air Quality (CMAQ)) to simulate the $PM_{2.5}$ concentrations and decoupled direct method (DDM, a sensitivity analysis technique for computing sensitivity coefficients simultaneously while air pollutant concentrations are being computed) to provide added information on the impact of prescribed burning on the local air quality. Measurements from low-cost sensors provide fine-scale $PM_{2.5}$ concentrations for evaluating the simulated concentrations and to assess how well the model captures $PM_{2.5}$ concentrations at those scales.

4.2 Materials and methods

4.2.1 Study Area

The domain we focus on in this study includes Dougherty, Lee, and Worth Counties in southwestern Georgia with a total population of about 150,000 (1.5% of state total population) and a total area of 330,000 ha (2.14% of state area). The per capita incomes for those counties are $19,210, $23,867 and $18,348 respectively, lower than the national per capita income ($27,334) according to 2015 U.S. Census data. Southwestern Georgia is one of the most active prescribed burning areas in the U.S.(Huang et al. 2018b). The annual total burned area of those three counties was around 30,000 ha in 2016 (5% of state total) (Huang et al. 2018b).

Three low-cost sensors (Plantower PMS 3003, Figure 4-1) were initially deployed at three high schools (Dougherty, Lee, and Worth County High Schools (DCHS, LCHS,

WCHS)) to measure the local $PM_{2.5}$ concentrations starting in May 16, 2017 and a fourth low-cost sensor was placed next to the state monitoring site (Figure 4-1) on March 14, 2018. Turner Elementary site (ID 130950007) is a suburban site located at Albany, Georgia. It was established in 1991 as part of the Georgia Air Protection Branch Ambient Monitoring Program. It only measures $PM_{2.5}$ with a BAM.

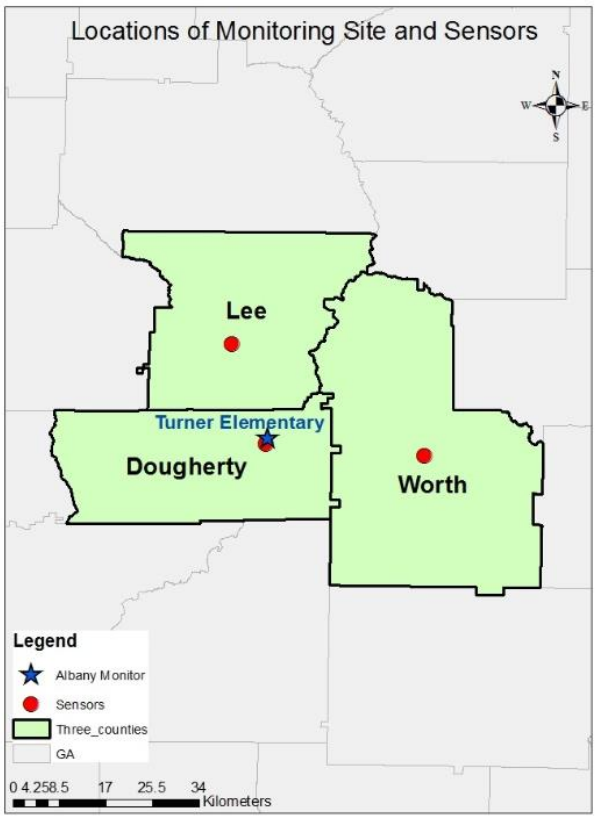

Figure 4-1 Locations of low-cost sensors and the Georgia Environmental Protection Division (EPD) monitor (*Albany in the figures represents the Georgia EPD monitor at Turner Elementary School.)

4.2.2 Low-cost Sensor Configuration

The low-cost PM sensor used for the field measurements in this study is Plantower PMS 3003. Kelly et al. (Kelly et al. 2017) conducted an evaluation of Plantower PMS 3003 in two locations associated with high levels of $PM_{2.5}$ ranging up to 700 $\mu g/m^3$: a controlled

wind-tunnel environment and an ambient environment. The study compared the low-cost sensor performance with research-grade, light-scattering instruments and found that the low-cost PM sensor correlates well with the instruments; however, this study also indicated that additional measurements under variable ambient conditions are needed. Zheng et al. (2018) evaluated the low-cost sensor (Plantower PMS 3003) in both low concentration suburban regions and a high concentration urban location. The measurements were compared against Federal Equivalent Methods (FEMs). The low-cost sensor performance was better when ambient $PM_{2.5}$ was higher.

4.2.2.1 Low-cost Sensor Calibration

The Plantower PMS 3003 low-cost sensor uses an optical method to measure PM_1, $PM_{2.5,}$ and PM_{10}. The light emitter is a He-Ne laser and a photodiode perpendicular to the light source is used to measure the scattered light. An internal program to the low-cost sensor calculates the concentration from the particle number concentration (number of pulses in the output waveform of the photodiode signal) and particle size (amplitude of the waveform of the photodiode output signal).

Time-weighted averaging from pre- and post-calibrations of the low-cost sensors at an urban background research site (Jefferson Street in Atlanta, Georgia, JST calibration) were applied to the raw data from the session to generate calibrated data. The calibrations were conducted by co-location with a regulatory instrument (tapered element oscillating microbalance (TEOM)). The inlet height of the low-cost sensor was approximately equal to the sampling inlet for the TEOM). A relative humidity correction factor (CF) of the form:

$$CF = a + b * \frac{RH^2}{1-RH} \quad (1)$$

where *a* and *b* are best-fit parameters, was used in the calibration and more details can be found in Zheng et al (2018).

There are limitations to using this calibration approach, including different light scattering properties of aerosols of different composition and size. Since light scattering differs by size and composition of particles it is best to calibrate the low-cost sensor in an environment similar to the one where it will be used, i.e., near wildland fire smoke. Sometimes, the RH correction calibration may not help. We tried to use the observation from EPD site as the referenced $PM_{2.5}$ and RH data from a nearby site (Southwest Georgia Regional Airport) in the National Weather Service (NWS) network to have another correction factor. However, there is no obvious correlation between the reference $PM_{2.5}$ and RH (Figure C - 1). Therefore, we decided to perform a linear regression calibration (Local calibration) between observations from the low-cost sensor and the BAM at EPD site Figure 4-2 and applied the relationship to all low-cost sensors' raw data.

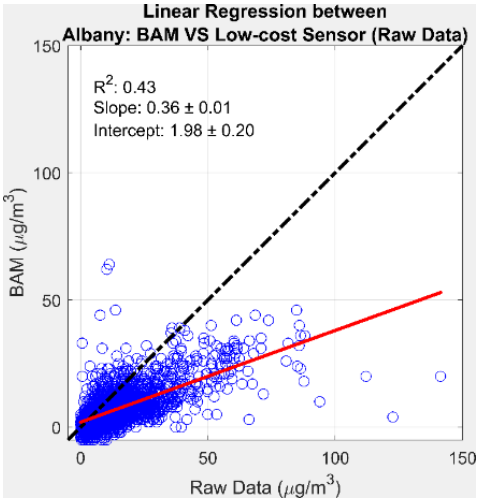

Figure 4-2 Comparison between BAM and low-cost sensor (raw data) collocated at the Albany EPD site

4.2.3 Air Quality Simulation

Air quality modeling using CMAQ was conducted for the period March 8 to March 15, 2018, we focus our analysis on a period capturing the exceedance day (daily $PM_{2.5}$ concentration > 35 $\mu g/m^3$): March 10, 2018. The daily total burn acres in the twelve counties surrounding the monitors on that day and the previous day are larger than 7,000 acres (Figure C - 2, Figure C - 3). Also, from March 13 to March 15, the daily total burn acres are larger than 5,000 acres every day. It is possible that those fires had a severe impact on air quality but this impact was missed by the observations both by the EPD monitor and the low-cost sensors. In order to evaluate the performance of low-cost sensors during the exceedance day and the feasibility of chemical transport models to capture the temporal and spatial variations of the fire impact, an eight-day period (March 08 – March 15) is simulated using Weather Research and Forecasting (WRF; version 3.6), a mesoscale numerical weather prediction model, and Community Multiscale Air Quality (CMAQ; version 5.0.2), a chemical transport model. CMAQ is equipped with Decoupled Direct Method (DDM), a sensitivity analysis technique for computing sensitivity coefficients simultaneously while air pollutant concentrations are being computed. CMAQ-DDM is used to calculate the impacts of prescribed burning emissions on $PM_{2.5}$. The prescribed burning emissions are calculated by using the BlueSky framework (Larkin et al. 2009) according to the burned area information from the Georgia Forestry Commission's (GFC) burn permit database. The fuel map in BlueSky is from the Fuel Characteristic Classification System (FCCS) (McKenzie et al. 2007). CONSUME model Version 3 (Joint Fire Science Program 2009) has been used to calculate fuel consumption in BlueSky. All fires are assumed to start at 11 a.m. and last 3 hours. The other emission data were

processed by Sparse Matrix Operator Kernel Emissions (SMOKE) based on 2011 NEI. The grid used by both WRF and CMAQ has a horizontal resolution of 4 km.

4.3 Results and Discussion

4.3.1 Comparison of the observations from low-cost sensors and the GA EPD monitor

The low-cost sensors were deployed and were found to have valid data at the four sites for periods of 3 to 12 months (Table 4-1).

Table 4-1 Low-cost sensor measurement dates

Low-cost Sensor Location	Measurement Dates
Dougherty County	May 16, 2017 – June 20, 2018
Worth County	May 16, 2017 – May 08, 2018
Lee County	May 16, 2017 – July 27, 2017; March 14, 2018 – April 20, 2018
Turner Elementary (EPD Albany site)	March 14, 2018- June 20, 2018

4.3.1.1 Comparison between low-cost sensors and BAM at the monitoring site using different calibration methods

Daily $PM_{2.5}$ observation from GA EPD's BAM at the Albany site and $PM_{2.5}$ concentrations show that the low-cost monitor captures the variations in the $PM_{2.5}$ levels and that the two different calibration methods (Local calibration and JST calibration) have a significant impact on performance (Figure 4-3). Daily $PM_{2.5}$ concentrations using JST calibration drift much more from BAM observations than those using local calibration. Comparison between the observations from the EPD's BAM observations at Albany site and daily $PM_{2.5}$ measurements by the low-cost sensor at DCHS (0.6 miles away from EPD

Albany site) show that local calibration performs better than JST calibration, with a higher R^2 and same slope (Figure 4-4). This illustrates that calibration at a distant site, even with an RH correction factor, may not improve the low-cost sensor performance due to the difference in the composition and level of $PM_{2.5}$. Because of the better performance obtained using local calibration, the following discussion will focus on results obtained using the local calibration only.

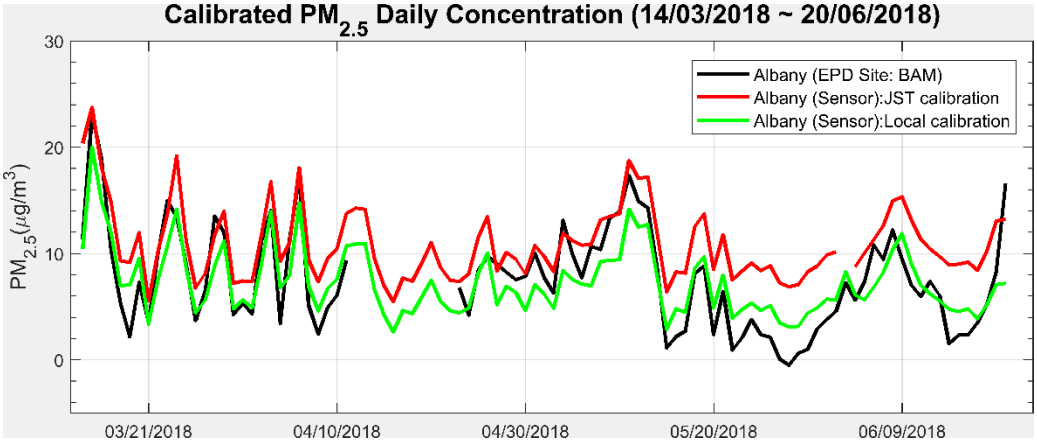

Figure 4-3 Daily $PM_{2.5}$ concentrations from low-cost sensors and BAM at GA EPD site (Albany) from March 14, 2018 to June 20, 2018 using different calibration methods

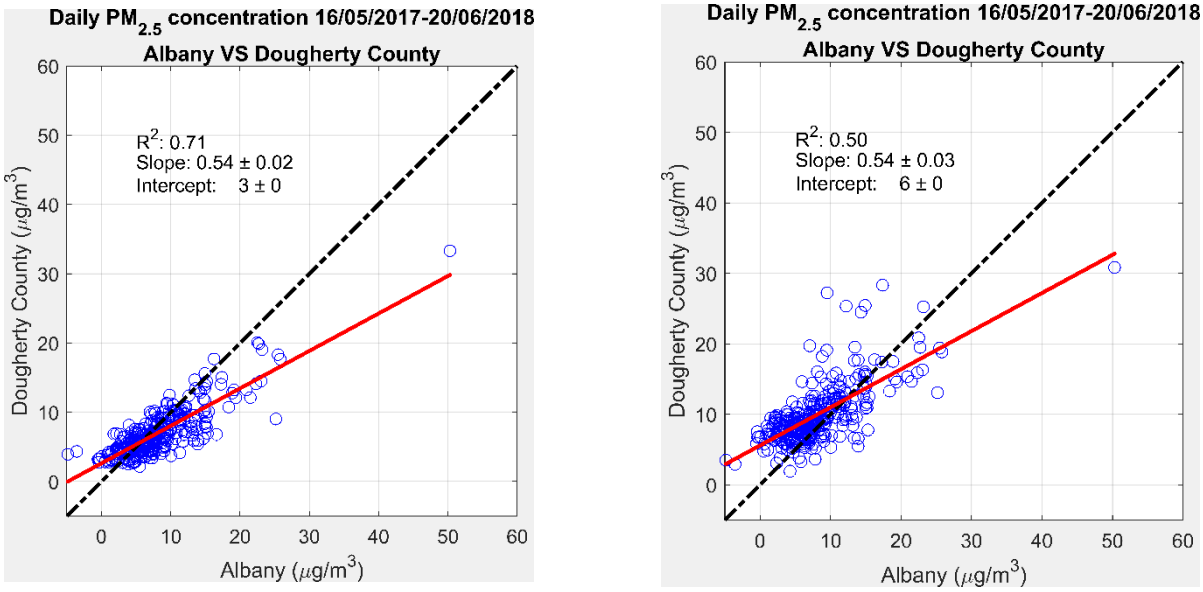

Figure 4-4 Comparison of the daily $PM_{2.5}$ concentrations between low-cost sensors and BAM at DCSH from May 16, 2017, to June 20, 2018: Local calibration (left), JST calibration (right)

4.3.1.2 Temporal trend of $PM_{2.5}$ from low-cost sensors and BAM at the monitoring site

Most of the hourly $PM_{2.5}$ concentration peaks detected by the BAM at the GA EPD's Albany site are captured by the low-cost sensors (Figure C - 4). However, low-cost sensors are less sensitive to changes in $PM_{2.5}$ concentrations than BAM due to detection limitations and sensitivity issue of the low-cost sensors, and record less variation (standard deviation: SD, of ~5 $\mu g/m^3$, Table 4-2) lower than that of BAM at the monitoring site (SD=8.68 $\mu g/m^3$). Differences in the observed peak levels between the sensors other than at the EPD site are likely due to the width of the fire plumes, leading to very heterogeneous levels on the scale of less than 1 km, and sometimes the plumes would miss some of the sensors completely. The Albany sensor observed very elevated levels from the plume on May 4and May 7, 2018, though the observation was less than that observed by the EPD BAM. On April 25, there are some high levels (>15 $\mu g/m^3$) detected by BAM and the Albany sensor also observed high concentration during the same periods, but the peak hour from the Albany low-cost sensor (9 a.m.) is one hour later than that from BAM (8 a.m.). The peak concentration from the low-cost sensor (33 $\mu g/m^3$) is higher than that from BAM (28 $\mu g/m^3$). The low-cost sensors at Albany and DCSH both capture a peak on April 26, however, there are no valid values on that day from BAM. Low-cost sensors can provide backup capabilities to regulatory monitors.

Daily trends of $PM_{2.5}$ concentrations from low-cost sensors and monitoring site show that the low-cost sensors did not capture levels as high as the EPD BAM (which recorded an exceedance) on March 10, 2018 (50 $\mu g/m^3$), but the concentration from DCHS, the closest low-cost sensor to the GA EPD site at Albany was quite elevated, (33 $\mu g/m^3$)

and is the highest value observed by the low-cost sensors during the measurement period (Figure 4-5).

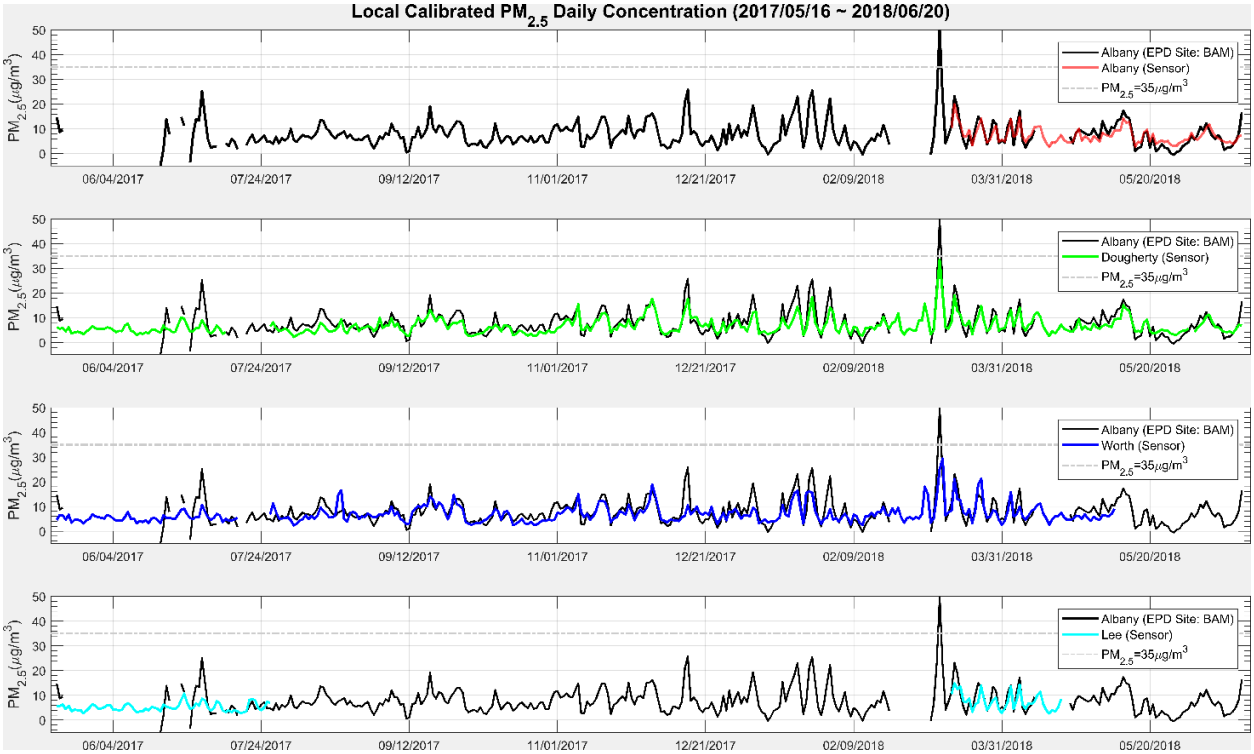

Figure 4-5 Daily $PM_{2.5}$ concentrations from low-cost sensors and BAM at GA EPD site (Albany) from May 16, 2017 to June 20, 2018

Table 4-2 Mean and standard deviation (Std) of $PM_{2.5}$ hourly concentrations from BAM at the GA EPD monitoring site in Albany and low-cost sensors

HOURLY	Albany (EPD)	Dougherty	Lee	Worth	Albany
Mean (µg/m^3)	7.98	6.77	6.00	6.95	7.28
Std (µg/m^3)	8.68	4.83	3.84	5.61	4.98

4.3.1.3 Comparisons of hourly and daily $PM_{2.5}$ concentration between low-cost sensors and BAM at the monitoring site

Both Figure C - 5 and Figure 4-6 demonstrate that low-cost sensors underestimate the $PM_{2.5}$ concentration with respect to the BAM observations at EPD's monitoring site with linear regressions slopes of less than 1. For DCHS, which is the closest low-cost

sensor to the monitoring site (0.6 miles), the R^2 (= 0.71) and slope (= 0.54) of the linear regressions of daily $PM_{2.5}$ concentration with the BAM observations at the monitoring site are the largest among the three low-cost sensors. Lower R^2 from Lee and Worth County comparisons do not mean those low-cost sensor performances are not good. With increasing distance between the low-cost sensors and EPD monitoring site, measured concentrations are more likely to reflect impacts from different fire plumes; therefore, those lower R^2 are expected. They show the spatial variation of fire impact on $PM_{2.5}$ concentrations and that a single monitoring site may miss high concentrations in the area.

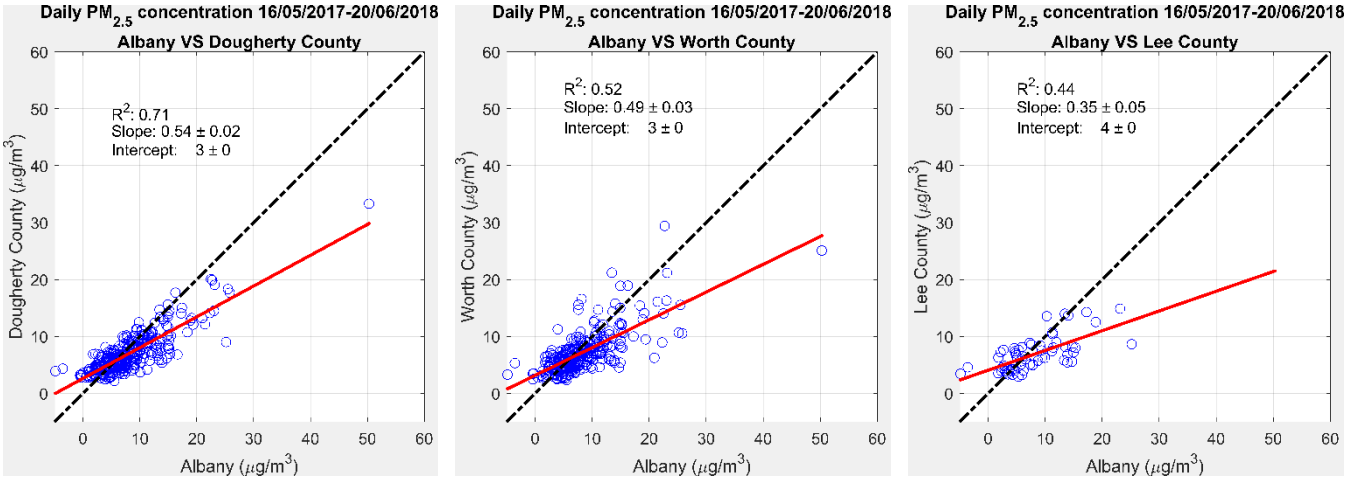

Figure 4-6 Comparison of daily $PM_{2.5}$ concentrations between high school low-cost sensors and BAM at GA EPD site (Albany) from May 16, 2017 to June 20, 2018 (LCHS sensor was down from July 28, 2017 to March 13, 2018 and April 21, 2018 to June 20, 2018; WCHS sensor was down from May 09, 2018 to June 20, 2018)

Both the hourly and daily $PM_{2.5}$ concentration regressions between BAM and low-cost sensor at Albany monitoring site have slopes less than 1 (Figure C - 6). The observation from the low-cost sensor of $PM_{2.5}$ concentration is higher than that from BAM below 10 µg/m^3 and lower for high concentrations. The comparisons between BAM with sensors at DCHS and WCHS with a cut-off at 95% of BAM observation (Figure C - 7)

shows that almost all the observations from DCHS and WCHS low-cost sensors are under the diagonal line. The R^2 (= 0.76) of daily comparison between BAM and low-cost sensor at Albany site is the largest one among the four low-cost sensors. The R^2 between BAM and the low-cost sensor at DCHS, WCHS, and LCHS are 0.78, 0.47 and 0.64 separately for the same period. The decreasing R^2 with increasing distance shows the spatial variation of fire impact that detected by different low-cost sensors but may be missed by the EPD site.

4.3.1.4 Intercomparison between low-cost sensors

The R^2 values for low-cost sensor intercomparisons are around 0.9 and with slopes very close to 1 (Figure C - 8). The increasing R^2 when comparing Worth, Lee and Dougherty low-cost sensors with the Albany low-cost sensor is a result of the spatial variation of the prescribed fire impact on $PM_{2.5}$ concentrations: WCHS, the farthest low-cost sensor from GA EPD site at Albany, has the lowest R^2 while DCHS, the closest low-cost sensor, has the highest R^2.

4.3.2 CMAQ-DDM results

The simulations do not capture the temporal variation of $PM_{2.5}$ concentrations well (Figure 4-7). The high $PM_{2.5}$ concentrations are mainly from fire impact (DDM results). The following discussion will focus on two periods: March 9 to March 10 and March 13 to March 14, 2018 to explain why the simulations are not correlated well with observations

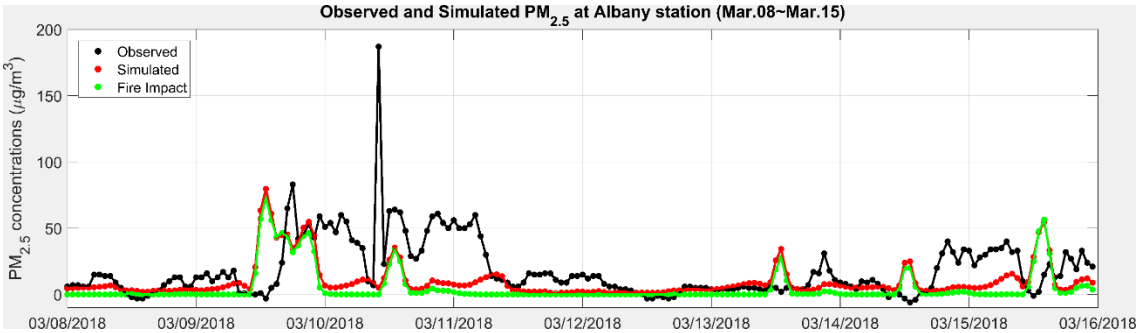

Figure 4-7 Comparison between observed and simulated $PM_{2.5}$ concentrations, and the fire impact at EPD site from March 8 to March 15, 2018

On March 9, the simulation predicts the peak concentration to be around 85 μg/m^3, which is in good agreement with the observed peak level, but to occur five hours earlier (Figure 4-8). This mismatch in the time of occurrence may be due to the actual start time and duration of the fires. All the fires were assumed to start at 11 a.m. and last 3 hours in the simulation; however, permit records for March 9 show fires approved to start in the afternoon and end at night. The actual start time and duration of fires are not known since no post-burn information is available. On March 10, the simulation does not capture the exceedance. The large difference between the observation and simulation at the beginning hours of March 10 may be mainly caused by meteorology. Observed wind speed is zero from 12 a.m. to 7 a.m. and 7 p.m. to midnight; however, WRF-simulated wind speed is larger than >2 m/s during the same periods (Figure 4-9). A systematic bias that leads to overestimated nighttime wind speed was reported in other applications of WRF (Garcia-Menendez et al. 2013; McNider et al. 2018). The observed peak at 10 a.m. may be caused by a new fire starting early and located close to the monitoring site. However, the

simulation does not have any fires starting before 11 a.m. and, since the boundary layer is changing rapidly at 10 a.m., it is also hard to tell which fire causes the peak.

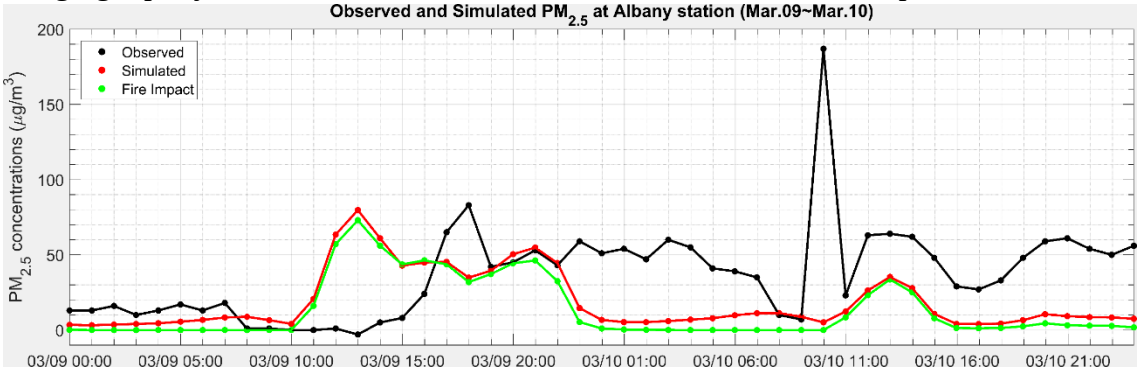

Figure 4-8 Comparison between observed and simulated $PM_{2.5}$ concentrations, and the fire impact from March 9 to March 10, 2018 at GA EPD site at Albany

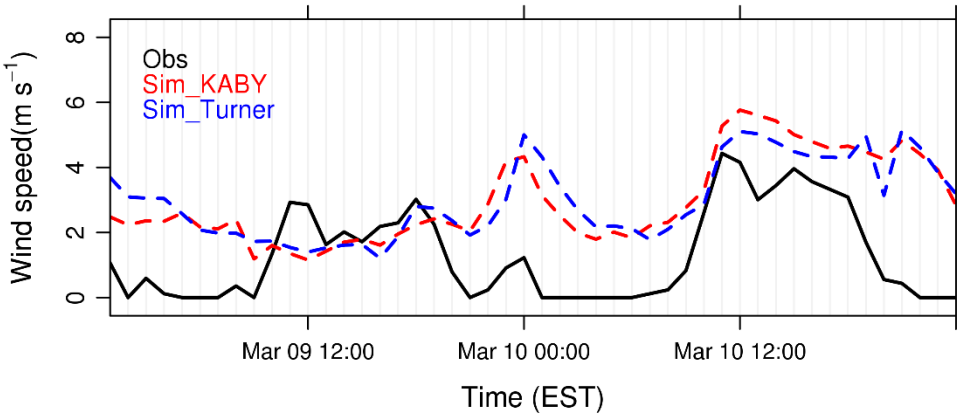

Figure 4-9 Wind speed from observations (Southwest Georgia Reginal Airport Site: KABY) and simulations

The DCHS and GA EPD site are located in two neighboring grid cells in the model simulations. At 6 p.m. on March 9, the simulated concentrations at these two grid cells are almost the same (~ 30 μg/m^3). However, the observations from the EPD site (80 μg/m^3) and low-cost sensor (40 μg/m^3) are quite different. The simulation does not capture this large spatial gradient in $PM_{2.5}$ concentrations (Figure 4-8, Figure C - 9). On the other hand, the peak concentration at the EPD site is detected at 10 a.m., while the peak at DCHS occurs at 11 a.m. The transport of smoke plume is detected successfully by the low-cost sensor.

For the WCHS low-cost sensor, the peak at 1 p.m. (Figure C - 10) is caused by medium size fires (51 to 150 acres) to the northwest of the low-cost sensor, marked by the orange circles in Figure C - 3 (right). There are some large fires (> 150 acres) to the southwest; however, their smoke plumes move to the southeast under northwesterly winds and just miss the EPD site as well as the DCHS low-cost sensor (Figure C - 11) when the observations at both locations are quite flat at low concentrations from 11 a.m. to 4 p.m.. Meanwhile, the southern counties, including Dougherty, are severely affected by those large fires as shown in Figure C - 11, but neither the EPD site nor the low-cost sensor does not indicate any presence of high $PM_{2.5}$ concentrations in the area.

4.4 Conclusions

The sparsity of air quality monitoring sites limits our ability to understand air pollution dynamics, evaluate air quality models, assess potential health impacts of air pollution and identify the most effective strategies to improve air quality and protect public health. The shortage of monitoring sites in Southwestern Georgia is of particular importance because of the widespread use of prescribed fires and their influence on local air quality and public health. As shown here, low-cost PM sensors can be used to detect prescribed fire impacts and provide the spatial information that may even be missed by model simulations. Further, they can provide back –up for instances when regulatory monitors may fail. However, low-cost sensor calibration is important and still needs further investigation to answer the question of what the maximum distance of the reference monitor to the ultimate low-cost sensor location should be for a reliable calibration. Calibration at a distant site with a different mix and level of $PM_{2.5}$ may result in poor low-cost sensor performance. RH correction factor calibration methods from the literature may

not help. In our case, i.e., using BAM as a reference monitor, the local calibration method was preferred.

Because of the highly non-homogeneous distribution of $PM_{2.5}$ concentrations in Southwestern Georgia, particularly when fire plumes are present, spatial gradients cannot be captured even with 4-km resolution model simulations. 1-km (or finer) resolution together with better knowledge of start and end times of the burns are needed to improve simulations. However, the accuracy of the fire impact simulation is highly dependent on accurately modeling the meteorology. The systematic high bias of wind speed at nighttime in the WRF model makes it harder to capture the temporal variation and level of pollution. Uncertainties in wind speed and direction limit the accuracy of the simulations. We recommend fusing model simulations with observations from a dense network of low-cost sensors for accurate estimation of exposures to smoke from prescribed burning.

4.5 Acknowledgment

This publication was made possible in part by funding from the Joint Fire Science Program under grant number 16-1-08-1 and NASA Applied Sciences Program under grant number NNX16AQ29G. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the supporting agencies. Further, the U.S. Government does not endorse the purchase of any commercial products or services mentioned in the publication. We thank everyone who helped us with the deployment of our sensors, in particular:

- Mr. Jeffrey Ross and Ms Chelsea Cox of Dougherty Comprehensive High School
- Mr. Kevin Taylor and Mr. Jayme Jones of Lee County High School

- Ms. Melissa Edwards and Ms. Kim Perrin of Worth County High School
- Mr. Ernesto Rivera, Mr. Ken Buckley and Ms. DeAnna Oser of Georgia Department of Natural Resources, Environmental Protection Division
- Dr. Fernando Garcia Menendez and Ms. Sadia Afrin of North Carolina State University for geocoded Georgia permit records data
- Dr.Susan M. O'Neill of Pacific Northwest Research Station for providing BlueSky framework
- Dr. Momei Qin of EPA for providing WRF figure
- Mr. Raj Lal of Georgia Institute of Technology for help setting low-cost sensors.

CHAPTER 5. EXPOSURE FIELD MODELING USING AIR QUALITY MODEL-DATA FUSION METHODS, AND COMPARISON WITH SATELLITE AOD-DERIVED FIELDS: APPLICATION OVER NORTH CAROLINA, USA

As published in *Air Quality, Atmosphere and Health*

Abstract

In order to generate air pollutant exposure fields for health studies, a data fusion (DF) approach is developed that combines observations from ambient monitors and simulated data from the Community Multiscale Air Quality (CMAQ) model. These resulting fields capture the spatiotemporal information provided by the air quality model, as well as the finer temporal scale variations from the pollutant observations and decrease model biases. Here, the approach is applied to develop daily concentration fields for $PM_{2.5}$ total mass, five major particulate species (OC, EC, SO_4^{2-}, NO_3^-, and NH_4^+), and three gaseous pollutants (CO, NO_x, NO_2) from 2006 to 2008 over North Carolina (USA). Several data withholding methods are then conducted to evaluate the data fusion method and the results suggest that typical approaches may overestimate the ability of spatiotemporal estimation methods to capture pollutant concentrations in areas with limited or no monitors. The results show improvements in capturing spatial and temporal variability compared with CMAQ results. Evaluation tests for $PM_{2.5}$ led to an R^2 of 0.95 (no withholding) and 0.82 when using 10% random data withholding. If spatially-based data withholding is used, the R^2 is 0.73. Comparisons of DF-developed $PM_{2.5}$ total mass concentration with the

spatiotemporal fields derived from two other methods (both use satellite aerosol optical depth (AOD) data) find that, in this case, the data-fusion fields have slightly less overall error, with an RMSE of 1.28 ug/m^3 compared to 3.06 ug/m^3 (two-stage statistical model) and 2.74 (neural network-based hybrid model). Applying the Integrated Mobile Source Indicator (IMSI) method shows that the data fusion fields can be used to estimate mobile source impacts. Overall, the growing availability of chemically-detailed air quality model fields and the accuracy of the DF field, suggest that this approach is better able to provide spatiotemporal pollutant fields for gaseous and speciated particulate pollutants for health and planning studies.

5.1 Introduction

Exposure to fine particulate air pollution ($PM_{2.5}$) has been associated with increased morbidity and premature mortality, suggesting that sustained reductions in pollution exposure could result in improved health and increased life expectancy (Gilboa et al. 2005; Sarnat et al. 2005; Pope et al. 2009; Matte et al. 2009; Solomon et al. 2012; Hubbell 2012). Estimating population exposure to $PM_{2.5}$ has traditionally been done by assigning measurements of a central ground monitor to people living within the region (Kanaroglou et al. 2005; Sampson et al. 2013). However, a number of studies have shown the limitations of using central ground monitor data as the exposure metric (Lefohn et al. 1987; Wade et al. 2006; Beelen et al. 2009; Kim et al. 2014; Dionisio et al. 2016). These limitations include monitoring sites in national regulatory networks that are relatively sparse across broad regions of the country (Hu et al. 2014a) and pollutant concentrations that can be impacted by local emissions, leading to local variations (Hu et al. 2014b). A variety of modeling approaches are now being used to better estimate pollutant concentration

variations not captured by monitors (Marmur et al. 2005; Johnson et al. 2010; Liu et al. 2012).

One approach to develop air quality fields is using chemical transport models (CTMs) that account for local variations affected by emissions and meteorology (Godowitch et al. 2015; Kim et al. 2015; Pleim et al. 2016). The Community Multiscale Air Quality (CMAQ (Binkowski 2003; Byun and Schere 2006b)) model is a state-of-the-science chemical transport model (CTM) designed to follow the dynamics of air pollutants from emissions. CMAQ captures spatial and temporal variations (Friberg et al. 2016), but is subject to errors due to limitations in insufficient characterization of meteorological (Yu et al. 2012) and emission inputs (Gilliland et al. 2008; Xiao et al. 2010; Ivey et al. 2015), as well as physical and chemical processes (Carlton et al. 2008; Tang et al. 2011; Ivey et al. 2016).

The objective of this research is to use the data fusion (DF) approach to develop spatiotemporal concentration fields for $PM_{2.5}$ mass, five PM species, and three gases for the state of North Carolina to support the University of North Carolina at Chapel Hill's health analysis of coronary heart disease patients in NC (McGuinn et al. 2017). The data fusion approach is developed at a spatial resolution of 12 km that combines observations from ambient monitors and data from CMAQ to better estimate ground-level air pollutant concentration fields for improved exposure estimates (Friberg et al. 2016). Several data withholding methods, which involve the use of monitor observations, were used to evaluate the stability of the data fusion method. A comparison of total $PM_{2.5}$ mass concentration is made between the results using unadjusted CMAQ pollutant fields, the data fusion application, ordinary kriging and two satellite aerosol optical depth (AOD) data-included

methods (Hu et al. 2014a; Di et al. 2016). These were compared as a part of evaluating the performance of various $PM_{2.5}$ exposure methods. Exposure fields of five PM species, and three gases were also compared between CMAQ results and data fusion method results.

5.2 Methods

Four statistical methods were used to create the spatiotemporal fields and the results were compared with each other and evaluated against observations. The first statistical method used was the data fusion method. The data fusion method combines observations and modeled pollutant fields, and was used during 2006-2008 period over North Carolina. (The data fusion method was actually applied from 2002 to 2010. 2006-2008 is in the middle part of that period and could be representative of the meteorological conditions experienced over that time.) The second and third methods were a two-stage statistical model and a neural network-based hybrid model, which both use satellite aerosol optical depth (AOD) and other data to develop $PM_{2.5}$ fields separately. Reliance on AOD data led to those methods being applied just to $PM_{2.5}$ mass, not individual PM or gaseous species. The fourth method uses ordinary kriging of observations at monitoring sites and was applied to develop $PM_{2.5}$ and CO fields. Other pollutants species were monitored at very few locations, limiting the amount of information available to develop spatio-temporal exposure fields as well as conduct a more thorough evaluation.

5.2.1 Air Quality Data

The observations used for data fusion come from the State and Local Air Monitoring Stations (SLAMS), Chemical Speciation Network (CSN) (Chu 2004) and Interagency Monitoring of Protected Visual Environments (IMPROVE) (Malm et al. 1994) networks. Observations from all available networks are utilized together. Pollutants include concentrations of three gases (carbon monoxide (CO), nitrogen dioxide (NO_2) and nitrogen oxide (NO_x)), $PM_{2.5}$ mass and five $PM_{2.5}$ components (elemental carbon (EC), organic carbon (OC), ammonium (NH_4^+), nitrate (NO_3^-) and sulfate (SO_4^{2-})) (Figure 5-1). Because of the limited number of monitoring sites for some species (e.g. CO, NO_2 and NO_x) in NC, we also included monitoring sites in neighboring states.

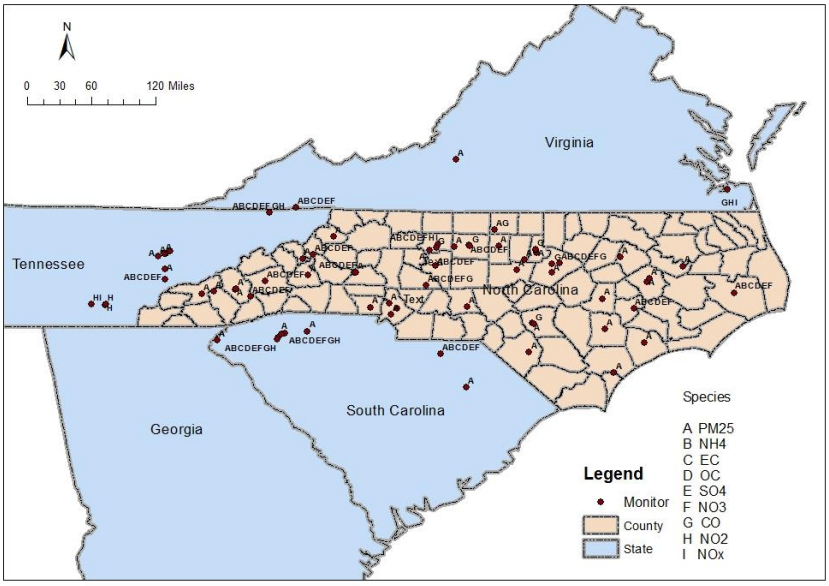

Figure 5-1 Ambient air quality monitor locations used in this analysis. (Not all monitor locations have all species.)

Twenty-four-hour average $PM_{2.5}$ concentrations for years 2006 to 2008 were collected from the EPA's Air Quality System Technology Transfer Network for use in the

two-stage statistical model. The MODIS aerosol data (collection 5) at 550 nm wavelength were obtained from the NASA Earth Observing System Data Gateway at the Goddard Space Flight Center.

5.2.2 Chemical Transport Model Simulated Concentrations

Pollutant concentration fields used in this paper are developed using CMAQ model version 4.5 at 12-km resolution for the 2006-2008 period over the North Carolina. A comprehensive model evaluation (Wyat Appel et al. 2008) of CMAQ version 4.5 conducted by the U.S. EPA showed that simulated particulate nitrate and ammonium are biased high in the fall due to an overestimation of seasonal ammonia emissions (Qin et al. 2015). The EPA evaluation also found that simulated carbonaceous aerosol concentrations are biased low during the late spring and summer due to the lack of some secondary organic aerosol (SOA) formation pathways in the model (Jathar et al. 2016; Woody et al. 2016).

5.2.3 Data Fusion

The approach used to combine the CMAQ-derived fields with observed pollutant concentrations was described in detail in Friberg et al. (Friberg et al. 2016). The method blends observations and CMAQ results based on spatial correlation analysis between observations and CMAQ simulations and generates a new field that captures local observations, as well as spatial variability from CMAQ. A summary is provided in the supplemental material.

Data fusion results were integrated with the Integrated Mobile Source Indicator (IMSI) method (Pachon et al. 2012) to estimate the influence of mobile sources on $PM_{2.5}$.

The IMSI method, which is developed for use in air quality and epidemiologic analyses, uses EC and NO_x as indicators of diesel vehicle (DV), and CO and NO_x as indicators of gasoline vehicle (GV) impacts. Here, the IMSI method, along with pollutant fields derived from the data fusion method, are used to provide spatiotemporal fields of mobile source impacts for use in source-specific, multipollutant, health analyses. The method is described in detail in the supplemental material.

5.2.4 Interpolation

Ordinary kriging (Cressie 1988) was applied to observed $PM_{2.5}$ and CO to develop air quality fields for comparison with the more advanced methods. $PM_{2.5}$ originates from multiple sources, both primary and secondary, whereas CO originates largely from mobile sources. $PM_{2.5}$ and CO are monitored at more sites than PM species and primary mobile source gases.

5.2.5 Methods Utilizing Satellite Aerosol Optical Depth for $PM_{2.5}$ Estimation

5.2.5.1 Two-stage Statistical Model

A two-stage statistical model (Hu et al. 2014a) employing satellite-retrieved aerosol optical depth (AOD) at 10 km resolution from Moderate Resolution Imaging SpectroRadiometer (MODIS) was used to develop $PM_{2.5}$ fields. The grids were restructured for comparison at 12 km resolution. The model includes a linear mixed effects module with day-specific random intercepts and slopes for AOD and meteorological fields as the first stage to account for the day-to-day variability in the $PM_{2.5}$-AOD relationship. The second

stage is a geographically weighted regression model to capture spatial variation. Details of the method are found elsewhere (Hu et al. 2014a).

5.2.5.2 Neural Network-based Hybrid Model

Di et al. (2016) applied another method that uses a neural network-based, hybrid model that includes satellite-based AOD data from MODIS, absorbing aerosol index (AAI), chemical transport model (GEOS-Chem) output, land-use terms, and meteorological variables. The method has been used to estimate the national $PM_{2.5}$ fields at 1 km × 1 km resolution. Detailed description is found in a previous publication (Di et al. 2016). We extracted the results for North Carolina for 2006 to 2008.

5.2.6 Model Evaluation Methods

The performance of the data fusion method was evaluated by using three data withholding methods, as described in following subsections.

5.2.6.1 Random Data Withholding

Ten groups of observational data were constructed, each group having 10% of the data randomly (not linked to specific monitors) withheld. Each group was run independently. Performance was assessed by comparing the simulated values to the data that were withheld for that iteration.

5.2.6.2 Randomly-based Monitor Data Withholding

Even though the random data withholding method is commonly used, it may overestimate the performance of the data fusion method. Monitor-based cross-validation

may better reflect performance of the data fusion method because it is representative of areas where no monitor is located as opposed to a situation where a measurement is missing. In this case, the entire set of 60 $PM_{2.5}$ monitors were randomly split into ten subsets with six monitors in each subset. For each of 10 cross-validation iterations, one subset (10% of monitors) was selected as the testing sample and the remaining nine subsets (90% of the monitors) were used to reapply the method. Estimates of the withheld monitor values were compared with the actual monitor values. This randomly-based monitor data withholding was repeated twice to check the stability of this evaluation to the random choice of monitor grouping. For NO_2 and CO, leave-one-monitor-out (LOO) was applied (i.e., in each test only one monitor data has been removed) due to the limited number of monitors available in the domain.

5.2.6.3 Spatially-based Monitor Data Withholding

Monitors may be clustered such that when one is removed there are nearby monitors that lead to the various methods being able to accurately estimate the pollutant levels for the removed monitor. This can result in an overestimation of a model's ability to provide accurate concentration estimates in a region with no monitors. Here, the entire set of monitors was spatially split into ten subsets (Figure D - 1) according to their locations, and withholding was performed with the spatially-based removed subsets.

5.3 Results and Discussion

5.3.1 CMAQ

As a baseline, the unadjusted CMAQ results are evaluated over the NC domain. Annual average $PM_{2.5}$, concentrations from CMAQ results (Table 5-1) are higher in 2007 for most species than in 2006 and 2008. For $PM_{2.5}$, the R^2 between pollutant observations and CMAQ simulations over the three-year period is 0.32 and a root mean square error (RMSE) is 5.16 ug/m^3. Linear regression (Figure 5-2, Table 5-2) between pollutant observations and CMAQ has a slope of 0.51. Evaluation results for other species tend to be have lower correlations (Table 5-2).

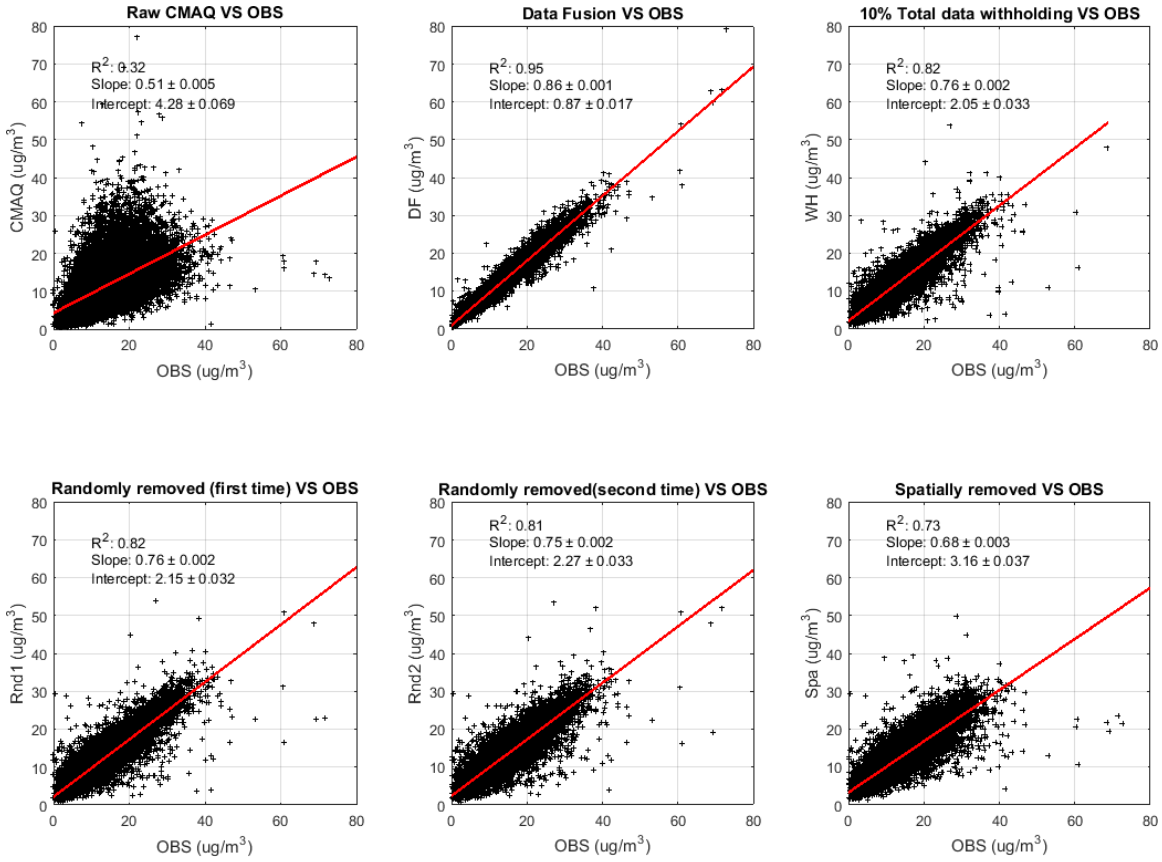

Figure 5-2 Linear regression between observation (OBS) and simulations ($PM_{2.5}$)

Table 5-1 Annual average concentrations from Data Fusion and CMAQ over the NC domain

Data Fusion	$PM_{2.5}$ (ug/m^3)	EC (ug/m^3)	OC (ug/m^3)	NH_4^+ (ug/m^3)	NO_3^- (ug/m^3)	SO_4^{2-} (ug/m^3)	NO_2 (ppb)	NO_x (ppb)	CO (ppb)
Monitor #	60	19	19	19	19	19	9	4	14
2006	11.12±5.09	0.47±0.44	2.03±1.82	1.39±0.75	2.16±2.91	3.91±2.40	9.00±4.30	8.39±14.26	302.14±104.53
2007	10.78±5.15	0.45±0.36	2.18±1.96	1.37±0.76	0.75±1.03	3.92±2.39	8.80±3.93	6.70±11.64	231.87±79.87
2008	9.70±4.69	0.31±0.24	1.87±1.81	1.44±0.78	0.47±0.69	3.31±1.85	8.25±3.79	5.28±7.50	279.83±92.97
CMAQ	$PM_{2.5}$ (ug/m^3)	EC (ug/m^3)	OC (ug/m^3)	NH_4^+ (ug/m^3)	NO_3^- (ug/m^3)	SO_4^{2-} (ug/m^3)	NO_2 (ppb)	NO_x (ppb)	CO (ppb)
2006	8.97±5.30	0.33±0.39	1.20±1.25	1.24±0.77	0.86±1.32	3.30±2.07	2.61±2.53	2.80±2.83	163.33±49.28
2007	9.09±5.62	0.31±0.27	1.51±1.54	1.17±0.73	0.96±1.50	2.79±1.71	3.16±3.06	3.16±3.54	165.48±52.60
2008	6.90±4.66	0.38±0.33	1.19±1.30	0.79±0.49	0.60±1.04	2.01±1.14	2.81±2.64	3.07±3.09	153.64±49.35

Table 5-2 Method Performance Evaluation (CMAQ, DF, DF-WH*) for $PM_{2.5}$ and $PM_{2.5}$ species (EC, OC, NH_4^+, NO_3^- and SO_4^{2-}) and mobile source related gases NO_2, NO_x and CO, 24-hour average values

SPECIES (# MONITORS)		PM_{25} (60) (ug/m^3)			EC (19) (ug/m^3)			OC (19) (ug/m^3)		
OBS VS Simulations		CMAQ	DF	DF-WH	CMAQ	DF	DF-WH	CMAQ	DF	DF-WH
BIAS (NME)		0.36	0.10	0.16	0.45	0.30	0.39	0.51	0.24	0.36
RMSE		5.16	1.28	2.48	0.51	0.35	0.36	2.03	1.06	1.27
R^2		0.32	0.95	0.82	0.28	0.62	0.32	0.26	0.71	0.45
Linear regression ($y = \alpha x + \beta$)	Slope(α)	0.507±0.011	0.864±0.003	0.76±0.002	0.651±0.004	0.811±0.003	0.44±0.01	0.562±0.003	0.892±0.002	0.59±0.01
	Intercept(β)	5.827±0.163	0.835±0.038	2.05±0.03	0.023±0.001	0.017±0	0.31±0.009	0.038±0.004	0.042±0.003	1.05±0.04
SPECIES (# MONITORS)		NH_4 (19) (ug/m^3)			NO_3 (19) (ug/m^3)			SO_4 (19) (ug/m^3)		
OBS VS Simulations		CMAQ	DF	DF-WH	CMAQ	DF	DF-WH	CMAQ	DF	DF-WH
BIAS (NME)		0.37	0.12	0.23	0.96	1.11	1.25	0.34	0.07	0.19
RMSE		0.74	0.25	0.42	1.15	1.59	1.49	1.86	0.42	0.98
R^2		0.34	0.92	0.67	0.41	0.52	0.40	0.60	0.97	0.82
Linear regression ($y = \alpha x + \beta$)	Slope(α)	0.851±0.003	1.029±0.001	0.72±0.01	1.306±0.008	1.881±0.01	1.62±0.03	0.699±0.002	0.9996±0	0.81±0.006
	Intercept(β)	0.019±0.002	0.005±0	0.51±0.02	0.013±0.003	0.008±0.004	0.25±0.03	0.05±0.004	0.0084±0.001	0.76±0.03
SPECIES (# MONITORS)		NO_2 (9) (ppb)			NOx (4) (ppb)			CO (14) (ppb)		
OBS VS Simulations		CMAQ	DF	DF-WH	CMAQ	DF	DF-WH	CMAQ	DF	DF-WH
BIAS (NME)		0.51	0.15	0.27	0.52	2.03	1.88	0.52	0.39	0.47
RMSE		7.14	2.44	3.16	12.29	49.44	33.17	268.76	231.16	178.04
R^2		0.18	0.81	0.78	0.26	0.76	0.37	0.08	0.26	0.24
Linear regression ($y = \alpha x + \beta$)	Slope(α)	0.573±0.008	0.878±0.003	0.61±0.007	0.496±0.008	3.11±0.022	1.87±0.11	0.368±0.003	0.82±0.005	0.25±0.008
	Intercept(β)	0.837±0.082	0.483±0.032	3.54±0.08	3.203±0.132	-2.025±0.378	13.36±2.24	63.445±1.087	91.417±1.91	362.75±3.49

* DF-WH: 10% Random data withholding; * NME: Normalized Mean Error.

5.3.2 Data Fusion

There are decreasing trends in annual average concentration for all species from 2006 to 2008 in the data fusion results (Table 5-1). The annual average concentrations for each species from the DF method are higher than those from the CMAQ results. The probability density distributions of all species concentrations are log-normally distributed (Figure D - 2).

Spatial plots of the annual averages for each of the nine pollutants show high concentrations in major urban centers (Figure 5-3; Figure D - 3; Figure D - 4). Emission impacts are evident near the major interstates in the NO_2, NO_x and CO fields. Concentrations at the western and eastern boundaries are much lower than the other areas because these are forest and coastal areas, respectively.

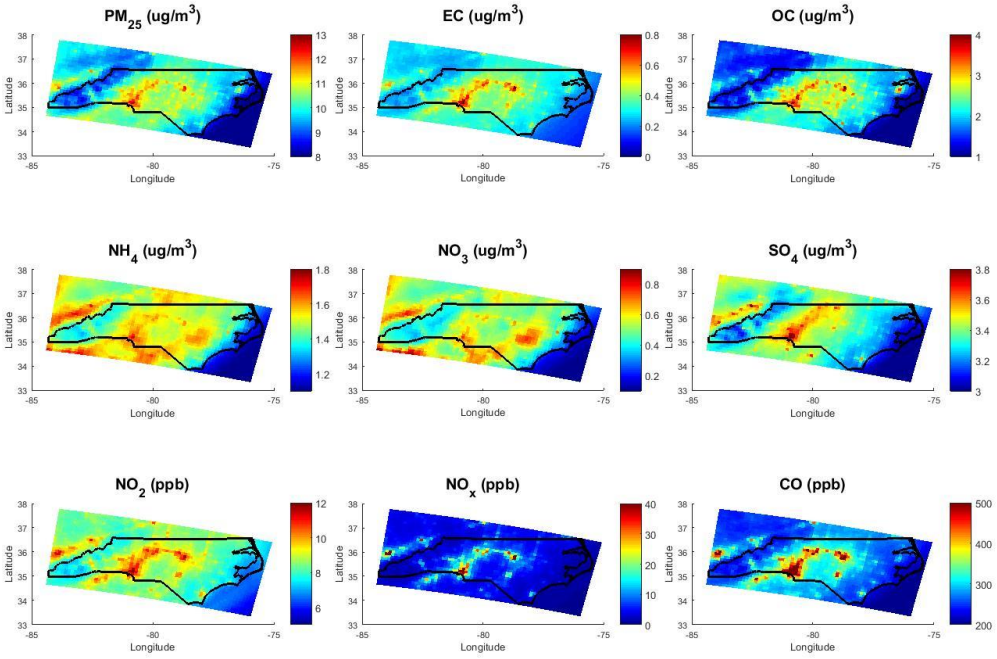

Figure 5-3 Annual average spatial distributions fields from data fusion, 2008

Monthly trends in North Carolina averaged over three years (Figure D - 5) show the concentrations of $PM_{2.5}$ and SO_4^{2-} are higher in the summer and lower in the winter in North Carolina, while NO_3^-, EC and OC are lower in the summer and higher in the winter. Concentrations of CO, NO_x and NO_2 are higher in the winter and lower in the summer. These trends are expected based on the atmospheric formation chemistry of the secondary components (i.e., sulfate formed in summer and nitrate in winter) and the mixing height (lower in winter) due to meteorological conditions.

Mobile source impacts are estimated using the IMSI method applied to the DF fields. IMSI impacts decrease in the summer and increase in the fall (Figure 5-4). The reduction of gasoline vehicle impacts is larger than the reduction of diesel vehicle impacts during the summer months. $IMSI_{GV}$ (emission-based IMSI value for gasoline) and $IMSI_{DV}$ (emission-based IMSI value for diesel vehicles) are higher in 2007 than 2006 and 2008 (Figure D - 6). The elevated impacts areas near highways indicate that the method captures mobile source activity and the data fusion fields are trustable (Figure D - 7).

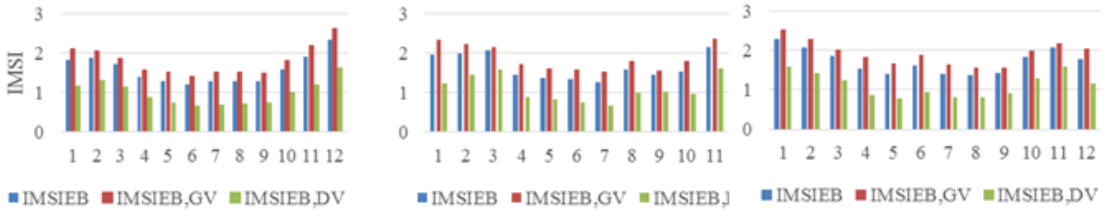

Figure 5-4 Monthly trends of $IMSI_{EB}$, $IMSI_{EB,GV}$ and $IMSI_{EB,DV}$ from 2006 to 2008 (unitless)

Temporal correlations between IMSI impacts and $PM_{2.5}$ concentrations indicate that highly populated and busy traffic areas have lower temporal correlations than other areas (Figure D - 8). The correlations between $PM_{2.5}$ and EC, CO and NO_x are low in rural areas (Figure D - 9). The low temporal correlation between $PM_{2.5}$ and the primary

pollutants is because much of the $PM_{2.5}$ in the area is secondary (Gertler et al.; Gertler 2005). The annual average spatial correlations between IMSI impacts and $PM_{2.5}$ concentrations are 0.72 (2006), 0.71 (2007) and 0.78 (2008).

Ten percent random data withholding (Figure D - 10) led to a R^2 of 0.82 (Figure 5-2) for $PM_{2.5}$, 0.24 (Figure D - 11) for CO and 0.78 (Figure D - 12) for NO_2. Reapplying the method led to very similar correlations (e.g., for $PM_{2.5}$, the R^2 was 0.81). Spatial 10% monitor withholding cross-validation (only applied to $PM_{2.5}$ due to the lack of monitors) led to a lower R^2 of 0.73 (Figure 5-2). The LOO results for CO and NO_2 also have lower R^2 values than the random data withholding, with a decrease from 0.24 to 0.10 for CO, and from 0.78 to 0.52 for NO_2. Although there is a small difference in $PM_{2.5}$ RMSE results of approximately 1.20 ug/m^3 between the 10% random data withholding results and the original DF data sets (Figure D - 13; Table 5-3), both of these values are much smaller than the CMAQ RMSE results of 5.16 ug/m^3. Spatial distributions of the maximum root-mean-squared-deviation (mRMSD: The maximum daily root-mean-squared-deviation value throughout the whole year.) for $PM_{2.5}$ show that the largest mRMSD are lower than 2, except in northeastern N.C. in 2008 (Figure D - 14; Figure D - 15). The RMSD of spatially-removed groupings (Figure D - 16) is similar to randomly-removed groupings (Figure D - 13; Figure D - 14) for $PM_{2.5}$, except for the northeast area of North Carolina in 2008 because of the limited monitors in this area (Figure 5-1). NO_2 results are similar, RMSE decreases from 7.1 ppb (CMAQ) to 2.4 ppb (data fusion) (Table 5-5). For CO, RMSE decreases from 269 ppb (CMAQ) to 231ppb (data fusion) (Table 5-4). RMSEs of LOO results for NO_2 and CO also show larger increases compared to 10% random data withholding results (Table 5-4, Table 5-5). All monitor-based withholding cross-validation

for $PM_{2.5}$, CO and NO_2 have larger RMSE and smaller R^2 than 10% random data withholding results.

The spatial 10% monitor withholding leads to a lower R^2 and higher RMSE for $PM_{2.5}$ as compared to random 10% monitors withholding (Table 5-3) with RMSE increases from 2.48 (random) ug/m^3 to 2.81 (spatial) ug/m^3. When removing values in spatially-similar groupings, kriging results are minimally impacted by distant observations. As a result, the CMAQ simulations are more heavily weighted and the performance of the withheld data fusion results worsens. The LOO test for NO_2 and CO show the influence of the distribution and quantity of the monitoring sites. CO monitors are located mainly in urban areas, while NO_2 monitors are distributed more widely. There are fewer monitors for both NO_2 and CO than for $PM_{2.5}$.

Table 5-3 Performance Evaluation for observation (OBS) and simulations ($PM_{2.5}$) using data withholding approaches, 24-hour average values

	Mean (ug/m^3)	Median (ug/m^3)	RMSE	NME*	R^2
Observations	12.7	11.5	0	0	1
CMAQ	10.9	9.6	5.16	0.38	0.32
Data Fusion	11.8	10.8	1.28	0.10	0.95
DF-10 % Random data withholding	11.8	10.8	2.48	0.16	0.82
DF-Random 10% monitors withholding (First test)	12.0	11.2	2.37	0.16	0.82
DF-Random 10% monitors withholding (Second test)	12.3	11.4	2.49	0.17	0.81
DF-Spatial 10% monitors withholding	12.3	11.5	2.81	0.19	0.73
Interpolation(Ordinary Kriging)	12.7	11.5	0.67	0.02	0.99
Ordinary Kriging: Random 10% monitors withholding (First test)	12.8	11.6	2.64	0.13	0.83
Ordinary Kriging: Random 10% monitors withholding (Second test)	12.7	11.6	2.74	0.14	0.81
Ordinary Kriging: Spatial 10% monitors withholding	12.6	11.5	3.23	0.19	0.71
Two-stage statistical model (no withholding)	12.8	11.7	3.06	0.15	0.81
Neural Network-based Hybrid Model (no withholding)	12.2	11.0	2.74	0.15	0.82

* NME: Normalized Mean Error.

Table 5-4 Performance Evaluation for observation (OBS) and simulations (CO),24-hour average values

	Mean (ppb)	Median (ppb)	RMSE	NME*	R^2
OBS	388	342	0	0	1
CMAQ	242	221	269	0.52	0.08
Data Fusion	461	421	231	0.38	0.26
DF-10 % Random data withholding	464	431	178	0.47	0.24
DF-Leave One-monitor Out	461	426	260	0.48	0.10
Interpolation (Ordinary Kriging)	391	346	24	0.05	0.99
Ordinary Kriging: Leave One-monitor Out	394	355	164	0.45	0.13

* NME: Normalized Mean Error.

Table 5-5 Performance Evaluation for observation (OBS) and simulations (NO_2), 24-hour average values

	Mean (ppb)	Median (ppb)	RMSE	NME*	R^2
OBS	11.0	10.1	0.0	0	1
CMAQ	7.4	6.0	7.1	0.51	0.18
Data Fusion	10.3	9.2	2.4	0.15	0.81
DF-10 % Random data withholding	10.1	9.3	3.2	0.27	0.78
DF-Leave One-monitor Out	10.3	9.3	3.8	0.27	0.52

* NME: Normalized Mean Error.

5.3.3 Ordinary Kriging Interpolation

Annual average $PM_{2.5}$ and CO spatial plots from kriging are shown in supplemental material (Figure D - 17). Linear regression (Figure D - 18; Figure D - 19) between ordinary kriging and observations has the highest R^2 and slope among all the methods. RMSEs are also very small, which are 0.67 ug/m^3 and 24 ppb, separately. Such performance is expected when using the same data in the application because of the ordinary kriging method's mechanism, so monitor-based data withholding was performed for evaluation.

The performance using monitor-based withholding for ordinary kriging are similar to data fusion results. R^2 for monitor-based withholding is larger than 0.70. Results for CO are worse than the total data interpolation; R^2 decreases from 0.99 (ordinary kriging) to 0.13 (ordinary kriging LOO) (Figure D - 19).

5.3.4 Methods using satellite-retrieved AOD for $PM_{2.5}$

5.3.4.1 Two-stage Statistical Model

The R^2 between observation and two-stage statistical model results is 0.81 (Table 5-3) lower than data fusion results (0.95, Table 5-2). The RMSE of two-stage statistical model (3.06 ug/m^3) is better than CMAQ data RMSE of 5.16 ug/m^3 when comparing simulated results with observations. A 10-fold cross validation (random data withholding) shows the three-year averaged R^2 is 0.78 and the averaged RMSE is 3.06 from 2006 to 2008.

5.3.4.2 Neural Network-based Hybrid Model

The linear regression between neural network-based hybrid model results and pollutant observations has an R^2 of 0.82 (Figure D - 20). The annual average spatial distributions fields (Figure D - 21) show a decreasing trend for $PM_{2.5}$ concentration from 2006 to 2008. The fields show that the method is also good at capturing the spatial information that urban areas have a high $PM_{2.5}$ concentration and rural areas have a lower concentration.

5.3.5 Comparison between CMAQ and Data Fusion for all species

Correlations between 10% random data withholding results and observations are higher than CMAQ and observations (Figure D - 10; Figure D - 13; Table 5-2). R^2 values for $PM_{2.5}$, EC, OC, NH_4^+, NO_3^-, SO_4^{2-}, NO_2, NO_x, and CO between observations and data fusion simulations increase compared to the correlations between observations and CMAQ simulations. RMSEs decrease and R^2 increases for all the species except NO_3^- and NO_x. The R^2 between observation and 10% random data withholding for $PM_{2.5}$ is 0.82. SO_4^{2-} also performs very well with a R^2 value of 0.82. R^2 value between daily CMAQ and data fusion results for each grid over the whole year for 2008 show that the highest values correspond to the grids that are nearest to monitors for all pollutants (Figure 5-5). R^2 values decrease as the distance to monitors increase, which indicates the accuracy of this method increases with the number of monitors used because of the high dependency on the number and locations of monitors to perform the kriging step in the data fusion method.

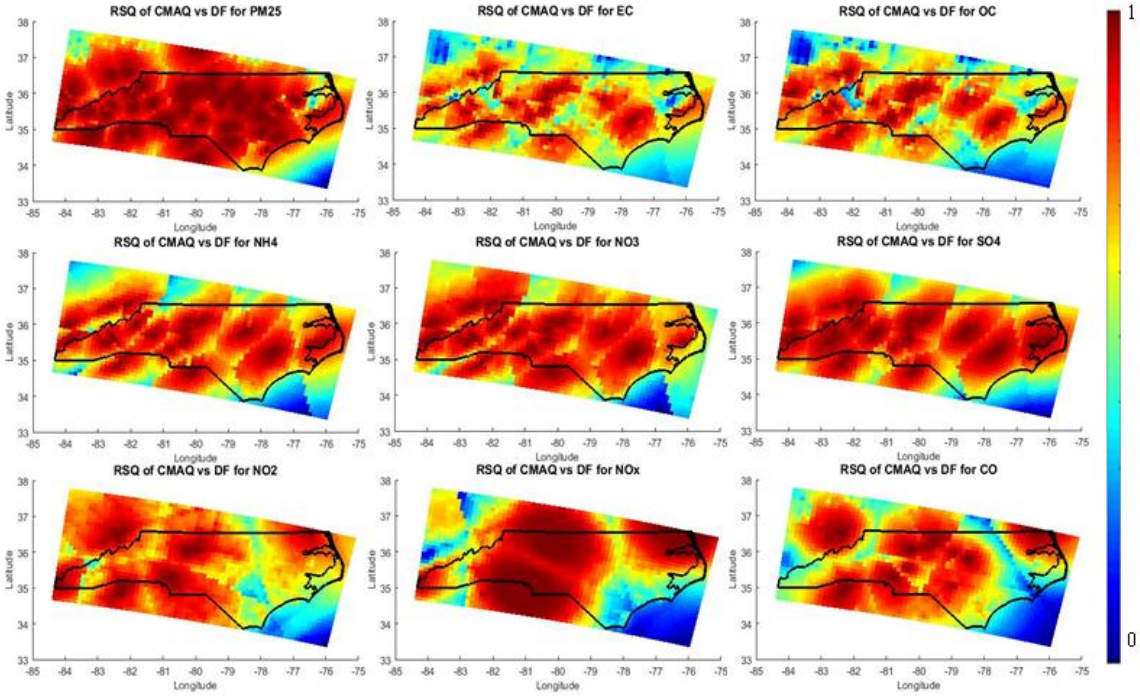

Figure 5-5 R^2 values of each grid for 2008

5.3.6 Comparison between Data Fusion and two-stage statistical model

The relationship between data fusion and two-stage statistical model results for $PM_{2.5}$ simulations during 2006 to 2008 are calculated using Deming regression (Deming 1943) to equally weight the two inputs because both data are estimated values from models (Figure 5-6). The grid-by-grid correlations over most of the domain have a value close to 1; however, the correlations in boundary areas are lower. Both the data fusion and two-stage statistical model capture the urban area $PM_{2.5}$ concentrations. Fewer monitors are located in the forested areas of NC, so the results from the two methods are not as strongly correlated. CMAQ secondary organic carbon formation is typical biased low in forested areas (Van Donkelaar et al. 2007; Zhang et al. 2007; Baek et al. 2011), which may contribute to low correlations with the two-stage statistical model. The two-stage statistical model can overestimate concentrations in the coastal areas of eastern NC (Figure D - 22)

because of the high relative humidity in the area, which leads to a bias in estimated $PM_{2.5}$ from satellite-retrieved AOD (Liu et al. 2005; Hu et al. 2013). The retrieval quality of the MODIS product is sensitive to vegetation cover and has difficulty distinguishing between the mixed land and water pixels, a limitation that might also contribute to the overestimation of the two-stage model along the coast. Lacking AOD data could be another limitation of these AOD-data included methods because of the satellite pattern and cloud cover days.

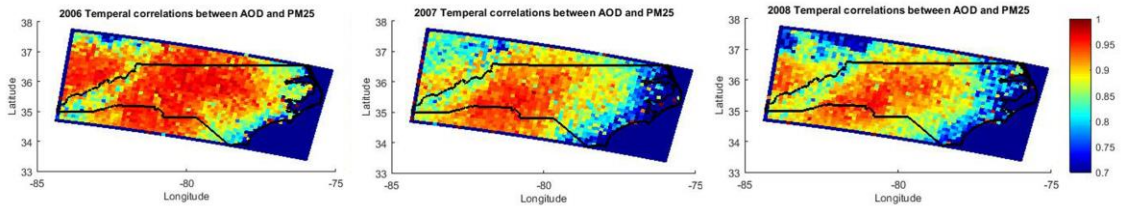

Figure 5-6 Temporal correlations (R) between data fusion and two-stage statistical model from 2006 to 2008

5.3.7 Comparison between Data Fusion and Hybrid model

Another comparison is made between the data fusion and Di et al. method (Di et al. 2016). Temporal Deming regression (Figure 5-7) shows the higher correlation in urban areas and lower correlation in the eastern, western boundaries and mid-south areas. This is similar to the comparison of data fusion and the two-stage statistical model results except in the mid-south areas, which is a national forest. The difference in annual average concentration in coastal areas (Figure D - 21, Figure D - 22, Figure D - 23) illustrate that the neural network-based hybrid model could provide more accurate spatial information because of the use of AAI and CTM outputs to improve accuracy.

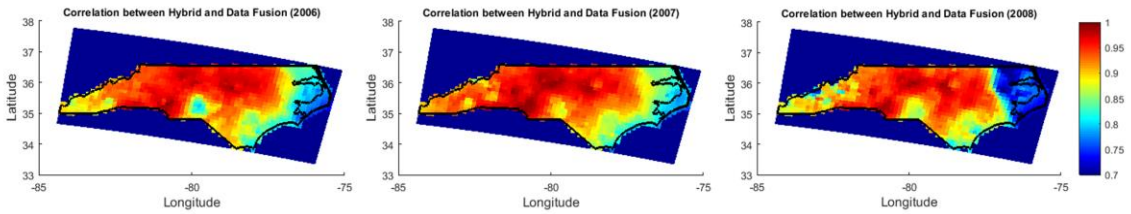

Figure 5-7 Temporal correlations (R) between data fusion and Harvard's hybrid method from 2006 to 2008

5.4 Conclusions

Application of the data fusion method for primary and secondary pollutants over North Carolina demonstrates that the method provides accurate concentration fields, especially for $PM_{2.5}$ total mass, OC, SO_4^{2-}, NH_4^+ and NO_2, capturing the spatial and temporal variations in both gaseous and speciated particulate matter concentrations. Capturing these variations is critical for improved estimation of exposures for health studies. Cross-validation with 10% random data withholding indicates that the DF results have little bias. CMAQ-modeled, non-data fused, concentration fields were subject to higher temporally and spatially varying bias and error, and lower correlations. These results demonstrate that the data fusion approach, as opposed to using CTM fields directly, should be used to provide spatiotemporal exposure fields for health studies that use daily air quality metrics. Using the DF method-derived fields to estimate mobile source impacts using the IMSI method also found that the results could be used in health studies.

This study also investigated the use of random data withholding versus withholding monitors randomly and based upon spatial clustering. Findings show that the data fusion method does provide accurate fields, but random data withholding may overestimate the ability of such methods to provide accurate concentration estimates in areas lacking monitors. The number and the distribution of monitoring sites affect the accuracy of the

data fusion method. The more widely the monitors are distributed, the more stable the data fusion method results. Observation availability is an important factor in the application and evaluation of the method according to some pollutants' performances such as CO, NO_2 and NO_x have very few monitors. Moreover, CO monitors are mainly located in urban areas. However, this research and previous studies demonstrate the benefits of the method versus the use of air quality model fields directly.

Spatiotemporal $PM_{2.5}$ fields derived using the CTM-based data fusion method compared well to similar fields derived using AOD and another chemical transport model. These and prior results suggest that the data fusion method provides a promising approach to develop exposure fields for health analysis across both urban and regional scales. A major advantage of CTM-based data fusion methods (which could potentially include the hybrid approach) over methods relying mostly on AOD to provide spatial variations is that it provides speciated $PM_{2.5}$ and gaseous pollutant fields.

5.5 Acknowledgements

We gratefully acknowledge the U.S. EPA, especially Valerie Garcia and K. Wyat Appel, for supplying CMAQ modeling results. The work of X. Hu and Y. Liu was supported by NASA Applied Sciences Program (grant numbers NNX11AI53G and NNX14AG01G, Principal Investigator: Liu). This publication was funded, in part, by U.S. EPA Grant Number R834799. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the U.S. government. Further, U.S. government does not endorse the purchase of any commercial products or services

mentioned in the publication. We also acknowledge the Southern Company and the Electric Power Research Institute (EPRI) for their support. We thank:

- Drs. Xinxin Zhai, Cesunica E. Ivey and Mariel D. Friberg for providing data fusion code
- Drs. Xuefei Hu and Drs. Yang Liu of Emory University for providing $PM_{2.5}$ exposure fields using two-stage statistical method
- Drs. Qian Di and Joel Schwartz of Harvard University for providing PM2.5 exposure fields using Neural Network-based Hybrid Model.

CHAPTER 6. THE IMPACTS OF PRESCRIBED BURNING ON AIR QUALITY AND HUMAN HEALTH: APPLICATION TO GEORGIA, USA

Abstract

Short-term exposure to fire smoke, especially $PM_{2.5}$, is associated with adverse health effects. In order to quantify the impact of prescribed burning on human health, a general health impact function has been used. A data fusion method has been applied to generate the exposure fields to $PM_{2.5}$ from prescribed burning during the burn seasons from 2015 to 2018. A method has been provided to distinguish the days and areas when and where prescribed burning has a major impact on local air quality for epidemiological studies to explore the relationship between prescribed burning and health effects. The results show a strong spatial and temporal variation of prescribed burning impact on health. Although southwestern, central, and east-central Georgia have large health impact rates, the number of Emergency Room (ER) visits related to asthma is small compared to metropolitan areas. Metro Atlanta is the most fire impacted area with largest number of ER visits for asthma due to fire impacts in Georgia.

6.1 Introduction

Epidemiological studies have shown the associations between short-term $PM_{2.5}$ exposure from fires and health endpoints such as mortality, respiratory effects, and cardiovascular effects (Rappold et al. 2011; Dohrenwend et al. 2013; Johnston et al. 2014; Faustini et al. 2015; Linares et al. 2015; Yao et al. 2016). However, most fire-related health

impact studies focus on wildfires. Prescribed burning, another type of biomass burning, which is a land management practice used to reduce wildfire risk in the U.S. also has large emissions. Such emissions remain one of the largest sources of $PM_{2.5}$ in the U.S. with an estimated 14% of total primary $PM_{2.5}$ emissions coming from prescribed burning, while wildfires account for 16% (US EPA 2014). This suggests that there is a tradeoff: land managers can use prescribed burning to reduce wildfires, and the related exposures, but should be mindful of the exposures resulting from prescribed burning as well. One advantage of using prescribed burning in this context is that they can be planned to minimize adverse human health impacts.

Georgia actively uses prescribed burning for land management with an annual statewide total burned area over one million acres (Huang et al. 2018b), one of the highest rates in the U.S., and an estimated 33% of the total $PM_{2.5}$ emissions come from prescribed burning (US EPA 2014). Epidemiological studies in Atlanta looking at the relationship between source specific $PM_{2.5}$ exposures and health effects like respiratory disease and cardiovascular disease (Sarnat et al. 2008; Darrow et al. 2014; Xiao et al. 2016; Krall et al. 2017) have found positive associations between same-day $PM_{2.5}$ concentrations attributed to primarily prescribed forest burning with cardiovascular disease-related Emergency Room (ER) visits. Krall et al. (2017) also found evidence of positive associations of respiratory disease ER visits with biomass burning $PM_{2.5}$.

In this paper, we use data fusion (Friberg et al. 2016), a method that merges air quality model simulations with observations from monitoring sites, to provide $PM_{2.5}$ exposure fields in health impact calculations. The Community Multiscale Air Quality (CMAQ) model (Byun and Schere 2006a), a chemical transport model, with the Decoupled

Direct Method (DDM) for source-specific impact estimation(Dunker 1984) is used to generate burn impacts from prescribed burning as a model input. A general health impact function is utilized for health impact assessment. The health impact from prescribed burning is quantified for GA at a spatial resolution of 4 km for the first four months of each year from 2015 to 2018, which are the most active burn months.

6.2 Materials and methods

6.2.1 Burn Impact Exposure fields

Daily total $PM_{2.5}$ concentrations at 4 km spatial resolution from 2015 to 2018 for the first four months in Georgia is estimated using the approach developed by Friberg et al. (Friberg et al. 2016; Huang et al. 2018a) that fuses observations from ambient monitors (Figure 6-1) and simulated pollutant concentrations from the Community Multiscale Air Quality (CMAQ, v5.0.2) model. The resulting fields capture the spatiotemporal information provided by the air quality model, as well as the temporal variations from the pollutant observations. This decreases model biases and errors. We applied CMAQ-DDM (Napelenok et al. 2006) to quantify air quality impacts associated with prescribed burning. The ratio of burn impact to total $PM_{2.5}$ for each day and each grid is applied to data fused total $PM_{2.5}$ fields to generate adjusted burn impact on $PM_{2.5}$. The adjusted burn impact will be compared with monitoring observations using the ratio between adjusted burn impact over the observation. The fire emissions are developed using the Bluesky framework (Larkin et al. 2009) with burned area information from the Georgia Forestry Commission's (GFC) burn permit database. All fires are assumed to start at 10 a.m. local time and last 6 hours according to permit records. Emissions data for other sources such as mobile,

agriculture, and biogenic emissions were projected from the 2011 NEI to the application year.

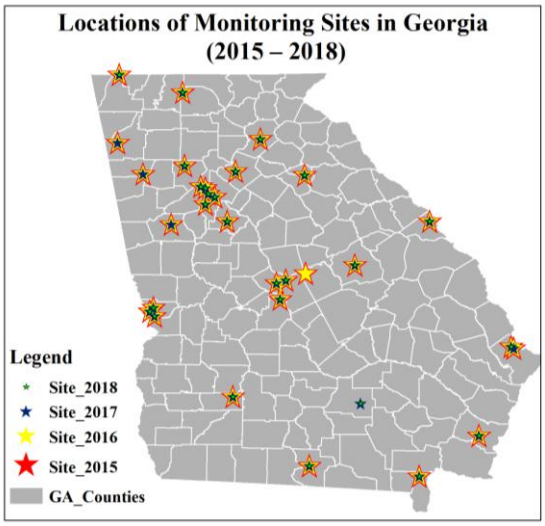

Figure 6-1 Locations of monitoring sites in Georgia: 2015 (Red stars), 2016 (Yellow stars), 2017 (Blue stars), and 2018 (Light green stars)

6.2.2 Health impact function

We used a log-linear relationship between air pollutant concentration change and health incidence to quantify the health impact from prescribed burning as follows:

$$\Delta Y = Y_0(1 - e^{-\beta \Delta PM}) \times Pop \quad (1)$$

where Y_0 is the baseline incidence rate for the health endpoint, β is the health effect estimate from the epidemiological study, ΔPM is the change in air pollutant concentration, and Pop is the population exposed to the air pollution. In our case, we focus on the ER visits for asthma as the health endpoint. The annual asthma-related ER visit rate for 2013 of Georgia is treated with the national asthma-related ER visit in 2013 (a rate 625.6 per 100,000

person) and converted to a daily rate by constructing weights based on observed daily ER visits, asthma counts during 2013 (Figure 6-2), as follows:

$$weight(i) = \frac{ER\ visits, asthma\ count_i}{\sum_{i=1}^{365} ER\ visits, asthma\ count_i}$$

where $ER\ visits, asthma\ count_i$ is the number of ER visits due to asthma on day i. Daily ER visits due to asthma in the Atlanta area for the first four months in 2013 has an average of 94.8 with a standard deviation of 17.5, the asthma-related ER visit rate is about 669.8 per 100,000 person.

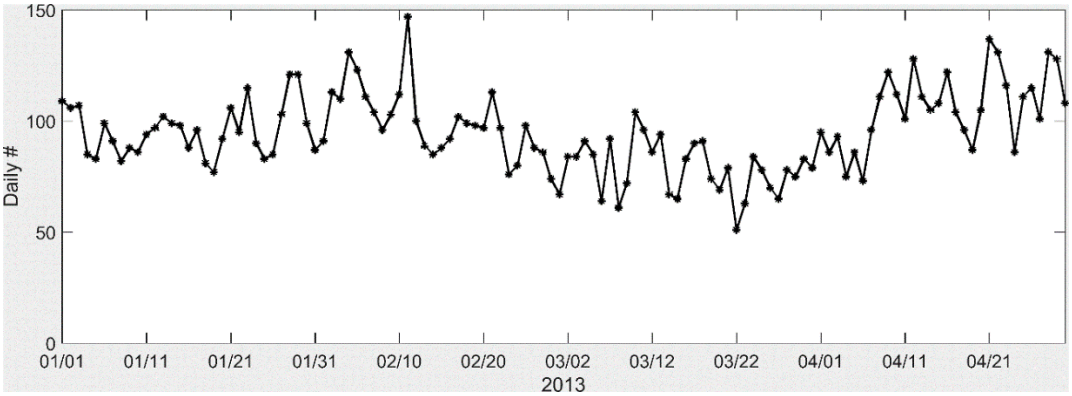

Figure 6-2 Observed daily ER visits, Asthma in Atlanta area (20 counties included): 2013

We used $\beta = 0.008$, which comes from a wildfire smoke exposure epidemiological study (Alman et al. 2016). ΔPM is the burn impact on total $PM_{2.5}$ after applying the data fusion method. We extracted the population from BenMAP-CE and allocated the 2010 block-level U.S. Census population to match the 4 km spatial resolution using PopGrid program provided by U.S. EPA (Figure 6-3) (US EPA). The health impact in following results part refers to the health impact of asthma-related ER visits.

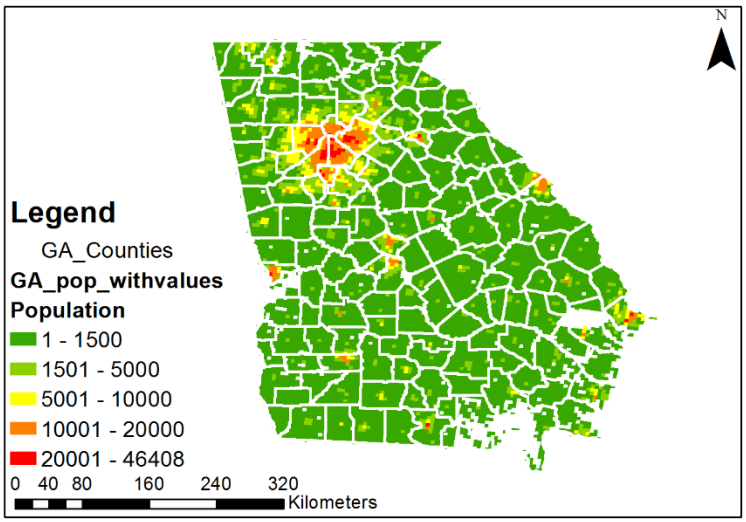

Figure 6-3 Georgia population (4 km resolution): 2010 U.S. Census (9,687,653)

6.3 Results and Discussion

6.3.1 Total $PM_{2.5}$ concentrations and fire impact exposure fields from CMAQ and data fusion (DF)

There's no obvious increasing or decreasing trend in average $PM_{2.5}$ concentrations from 2015 to 2018 during the prescribed burning season (Table 6-1). CMAQ simulations of total $PM_{2.5}$ concentrations were biased low compared to observations. Fire impact increases since 2016 are due to more area being burned over last three years (Figure 6-4) and the larger fuel loads and emissions caused by drought in the southeastern U.S. in the fall of 2016 (Park Williams et al. 2017). Spatial plots of the monthly averages for total $PM_{2.5}$ and fire impact (Figure E - 3 ~ Figure E - 6) show that high concentrations in the southwestern and east-central GA are mainly due to the active prescribed burning. January, February, and March are more active burn months than April. March is the most active burn month accounting for about 40% of the total burned area among the first four months

(Figure 6-4). 2018 has a total burned area over one million acres for the first four months which brings to the largest burn impact over the last four years (2015 – 2018).

Table 6-1 Average total $PM_{2.5}$ concentrations and burn impacts at monitoring sites during the first four months of 2015 – 2018

Mean(μg/m^3)	2015	2016	2017	2018
Observation	8.55±4.33	8.34±4.35	8.82±5.31	8.67±4.99
CMAQ	6.62±4.45	6.13±4.32	6.32±3.99	6.23±4.00
CMAQ_B*	0.83±2.82	0.76±2.96	0.87±1.62	0.97±2.26
Data Fusion	8.41±3.81	8.02±3.68	8.74±4.59	8.23±4.19
DF_B*	0.91±1.97	0.86±1.77	1.06±1.77	1.24±2.42

*_B: burn impact in $PM_{2.5}$

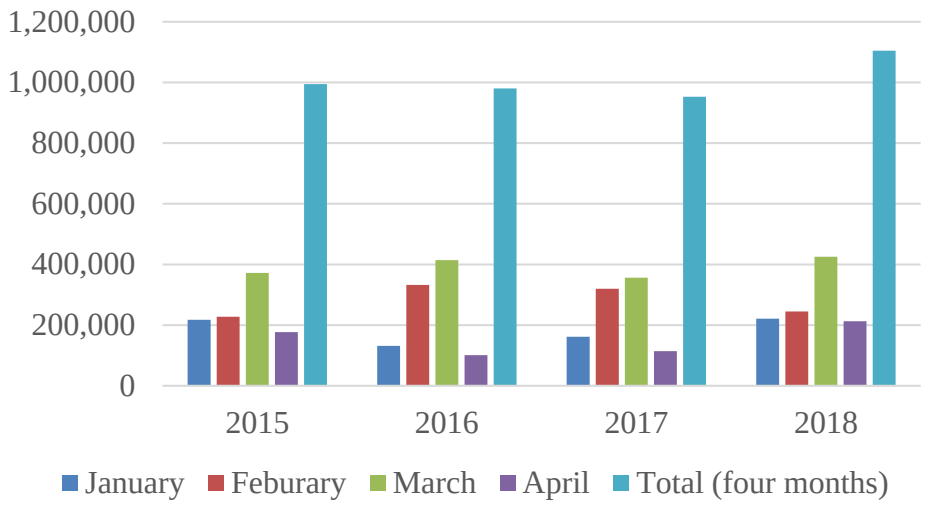

Figure 6-4 Monthly total burned area from 2015 to 2018, first four months

Comparisons between observations and CMAQ (Figure E - 1, Table 6-2) over four years of daily total $PM_{2.5}$ concentration during January to April have slopes less than 0.5 and R^2 range from 0.13 to 0.30. Based on the recommended performance statistics to assess photochemical model performance (24-hr $PM_{2.5}$ criteria: *R* (correlation coefficient) >0.4) from Emery et al. (2017), only the R^2 of 2016 does not meet the criteria. All years' NMEs meet the criteria (< 50%) and goal (< 35%). Here, the selection of statistical goals and

criteria was based on regional photochemical grid models applications being developed at the time to support U.S. regulatory actions for $PM_{2.5}$ and regional visibility (Emery et al. 2017).

Table 6-2 Normalized root mean square error (NRMSE) and normalized mean error (NME) of CMAQ simulated and data fused daily total $PM_{2.5}$ concentration with respect to observations during the first four months of each year (2015 – 2018)

	NRMSE				NME			
	2015	2016	2017	2018	2015	2016	2017	2018
CMAQ	0.57	0.64	0.59	0.62	0.23	0.26	0.28	0.28
DF	0.23	0.27	0.24	0.24	0.02	0.04	0.01	0.05

Comparisons between observation (OBS) and DF (Figure E - 2) show improvement after applying data fusion method with slopes closer to 1 and R^2 increasing to around 0.8. Data fusion NRMSEs decrease by ~60% compared to CMAQ NRMSEs. NMEs also decrease and are close to zero. All those results show better performance with the application of data fusion method. Occasionally simulated daily $PM_{2.5}$ concentrations from CMAQ are larger than the observations. These are due to simulated fire impacts that are not captured by the observations at the sparse monitoring sites in the region, leading to large differences between the data-fused and original CMAQ results (Figure E - 3 ~ Figure E - 6). This can lead to the data fusion process decreasing the modeled impact of fires. We also made a comparison between observed daily total $PM_{2.5}$ and ratio of adjusted burn impact to observed $PM_{2.5}$ (Figure 6-5 and Table 6-3). The grey dash lines represent 95 percentile of observation (vertical) and 30% ratio of fire impact to observed $PM_{2.5}$. The reason we chose 30% here is because 33% of $PM_{2.5}$ emissions come from prescribed burning in Georgia according to 2014 NEI. Red dots are the days with high observed concentrations due to fire impacts as determined using model results. Epidemiological

studies could focus on those days to find the relationship between short-term high-level $PM_{2.5}$ exposure due to fire impact and health effects over a series of single-day lags. Blue dots are the days that the fire impact is still the major source of total $PM_{2.5}$. However, due to the low observed $PM_{2.5}$ concentrations, those days may not arise people's attention but could cause health impact from exposure to fire smoke and are also worth investigating. Green dots are the days that are affected by fires but also other sources at the same time. Nearly 13% of observations in 2018 are dominated by prescribed fires, somewhat larger than previous years.

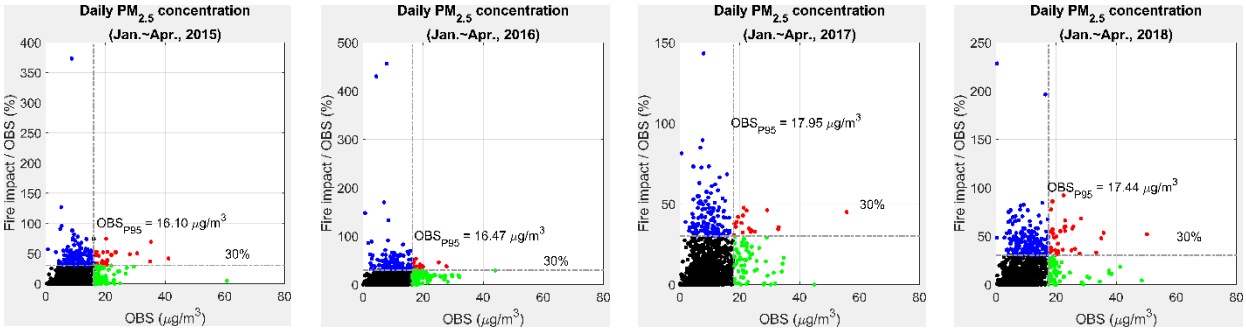

Figure 6-5 Comparison between daily total $PM_{2.5}$ observations (OBS) and ratios of adjusted fire impact to OBS from 2015 to 2018, first four months

Table 6-3 Numbers of days in each quadrant in Figure 6-5

	2015	2016	2017	2018
Total	2069	2136	1565	1762
High fire impact / High $PM_{2.5}$	27	14	18	36
High fire impact / Low $PM_{2.5}$	141	124	133	192
Low fire impact / High $PM_{2.5}$	75	93	60	52
Low fire impact / Low $PM_{2.5}$	1826	1905	1354	1482

6.3.2 Health impact from prescribed burning

Monthly total asthma-related ER visits health impacts from prescribed fire from 2015 to 2018 show spatial and temporal variation (Figure 6-6). Southwestern, central and east-central GA have large health impacts due to the intense prescribed burning activity. Macon Metropolitan Statistical Area (MSA) (Bibb County), Atlanta MSA (Fulton, Gwinnett, DeKalb and Cobb Counties), Albany MSA (Dougherty County), Augusta MSA (Columbia and Richmond County), Warner Robins MSA (Houston County), Valdosta MSA (Lowndes County), and Columbus MSA (Muscogee County) have larger health impacts in terms of absolute numbers due to both a large population and high level of fire impact. Although the prescribed burning does not have that much impact on air quality in Atlanta MSA (Figure E - 3 ~ Figure E - 6), the large population still leads to a large health impact. January, February, and March experience the larger health impacts due to more active prescribed burning (Figure 6-4). Typically, April is not an active burn month, however, 2018 had more burns in April than February in the central and east-central GA, and those affected Atlanta MSA's air quality. 2015 April also had more burns in southern GA, though larger impacts were limited to less populated regions.

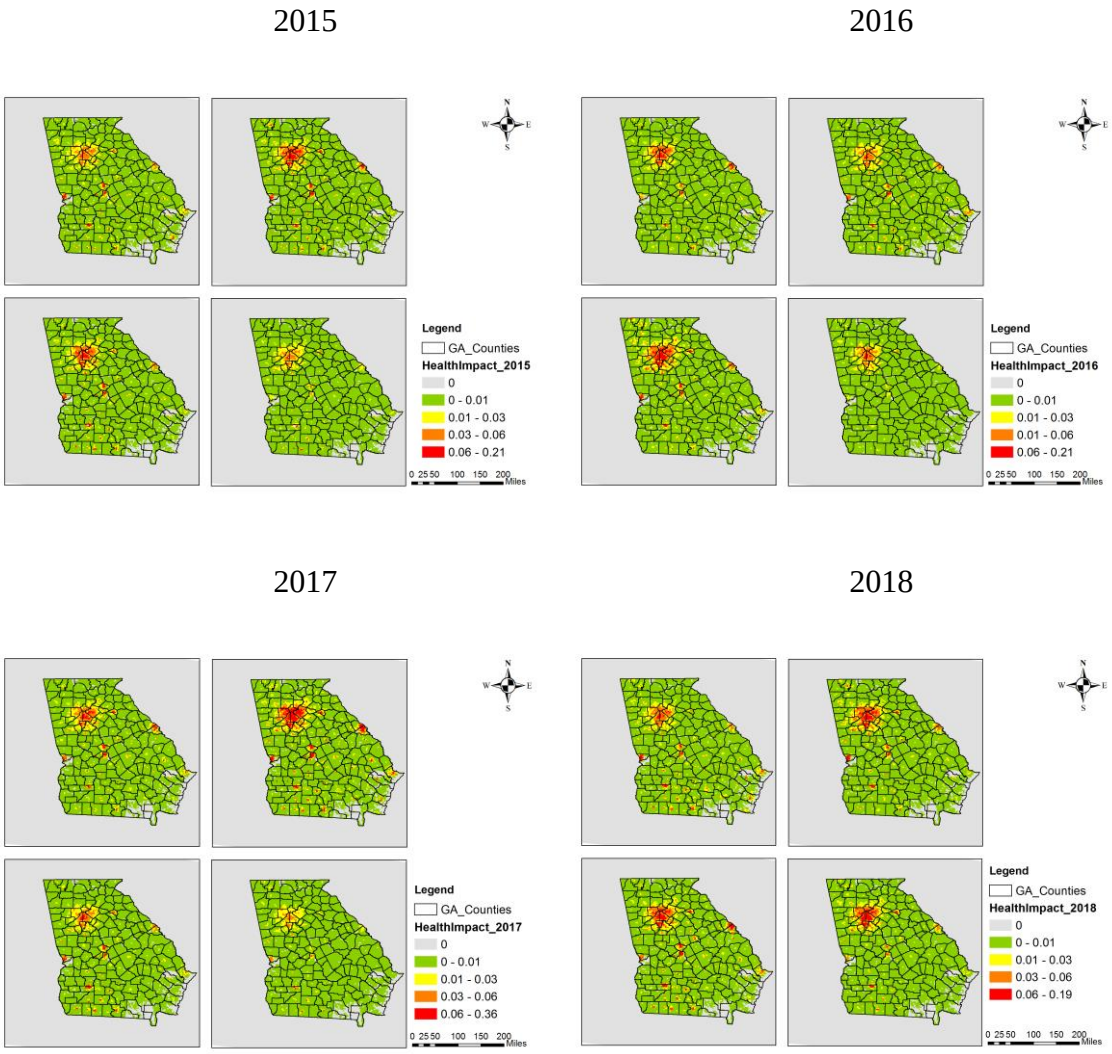

Figure 6-6 Monthly total asthma-related ER visits health impact from prescribed fire from 2015 to 2018, first four months

Daily asthma-related ER visits due to prescribed burning for each year (Figure 6-7) and four years together (Figure 6-8) for each of the four months separately show that February and March have larger health impact than January with higher daily health impact. The health impact from first three months has larger temporal variation over the years than April with larger interquartile range (IQR). There is a slightly increasing trend of daily average health impact from 2016 to 2018. The temporal variation of January among four years is similar to the four months' total temporal variation. February in 2017 has larger daily health impact compared to the other years due to larger emissions from drought season. 2018 April also has larger daily health impact with more burned areas compared to previous years indicate that the burn season may extend with the need to burn more areas.

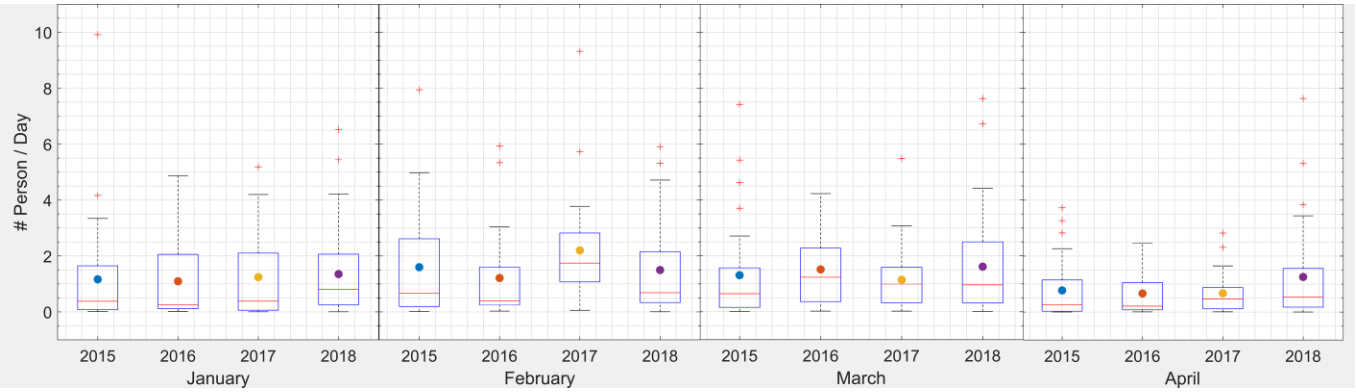

Figure 6-7 Daily asthma-related ER visits due to prescribed burning for each year (2015 – 2018), each month (January – April) The central mark indicates the median, the point indicates the mean, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively.

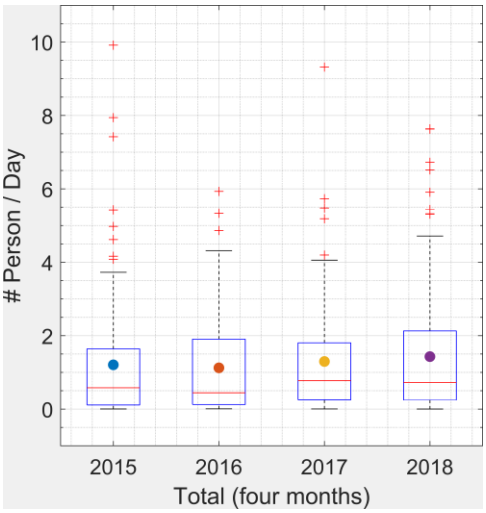

Figure 6-8 Daily asthma-related ER visits due to prescribed fire for 2015 to 2018, first four months. The central mark indicates the median, the point indicates the mean, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively.

The total health impacts increase from 2016 to 2018 (Figure 6-9) as the prescribed burning season appears to be getting longer (Figure 6-4). April used to be a less health impacted month compared to first three months, but in 2018 the total health impact of April doubled compared to previous years. Although, the burned area in 2015 and 2018 February is about 60% of that in March (Figure 6-4), the total health impact of those two months in each of the two years is similar (Figure 6-9). This result indicates that there are more populated areas (Atlanta MSA and Augusta MSA) affected by prescribed burning in February in those two years than March. 2017 February has the largest health impacts across the reporting years with about 62 ER visits due to asthma, a rate of 6.4 per 1,000,000 people. There is less difference in total health impact among different months in 2018 compared to previous years. There are about 145 ER visits due to asthma because of prescribed burning impact, a rate of 15 per 1,000,000 people in 2015 during the first four months. The number increases by about 18% in 2018 compared to 2015. Total number of

ER visits increases about 15% from 2016 to 2017, especially for February in 2017 with increases of about 77%. The number of 2018 April ER visits is 38, increases over 60% compared to 2015 (23) and 90% compared to 2016 (20) and 2017 (20).

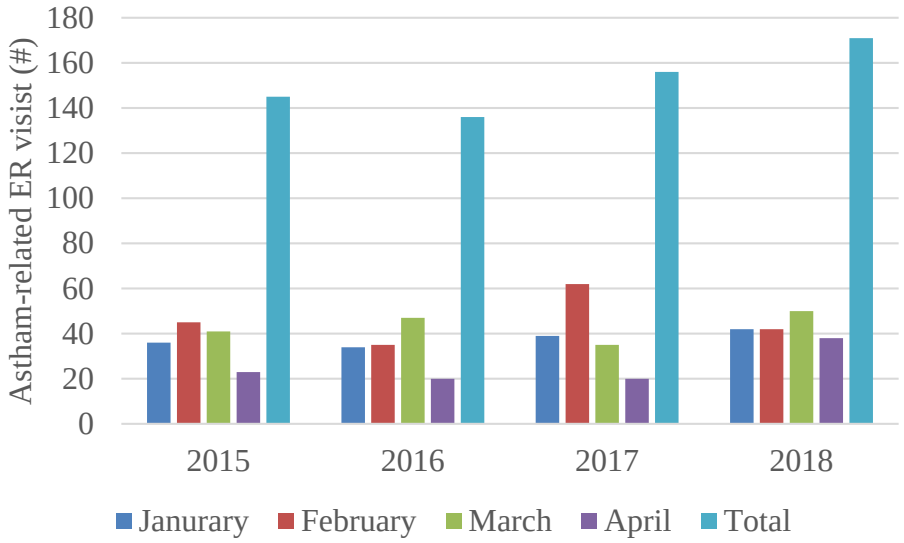

Figure 6-9 Total asthma-related ER visits from prescribed fire from 2015 to 2018, first four months in Georgia

Table 6-4 Monthly total ER visits due to asthma in Georgia

	January	February	March	April	Total
2015	36	45	41	23	145
2016	34	35	47	20	136
2017	39	62	35	20	156
2018	42	42	50	38	171

Fulton County has the largest health impact due to the largest population (Figure 6-10). There are 10 ER visits due to asthma because of prescribed burning impact in 2015 during the burn season. The number increases about 30% to 13 in 2018. Gwinnett, DeKalb and Cobb Counties, all within the Atlanta MSA, are the other three counties that have over 5 people visit ER because of asthma issue during the burn season every year (Figure 6-11).

The estimated asthma-related ER visits due to prescribed burning in Atlanta MSA has an average about 66 during the reporting years, which is about 0.58% compared to the observed asthma-related ER visits (11,372) in Atlanta MSA from 2013. Dougherty County has the largest health impact in the southwestern GA, but with an average about 4 people visit the ER related to asthma during the burn season for the four years (2015 – 2018) due to the small amount of population. The spatial distribution of asthma-related ER visits is similar to the observation from the Georgia Department of Public Health for 2014 (Figure E - 7). Atlanta MSA has the most amount of people visit the ER due to asthma caused by burn impact.

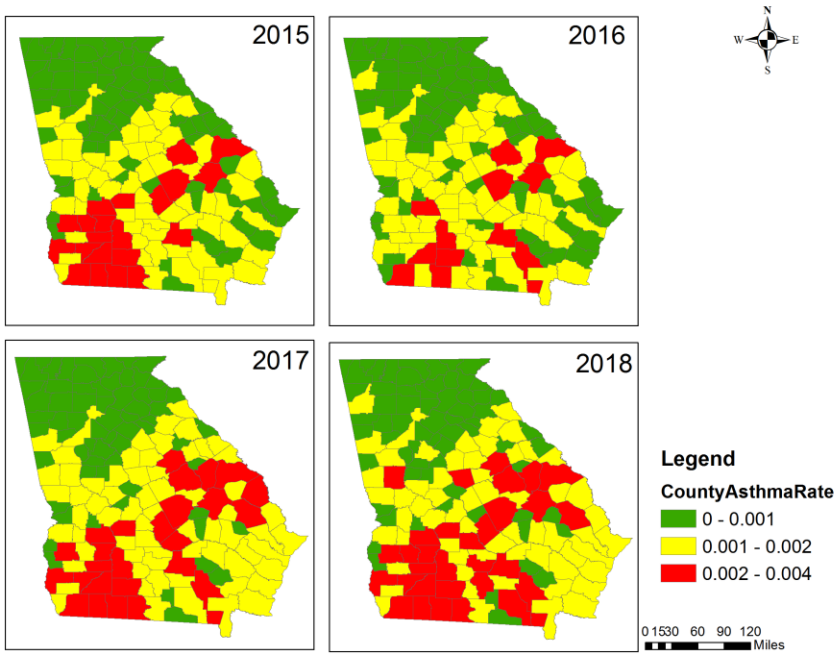

Figure 6-10 Asthma-related ER visits rate due to prescribed burning by county in Georgia from 2015 to 2018, first four months (per people)

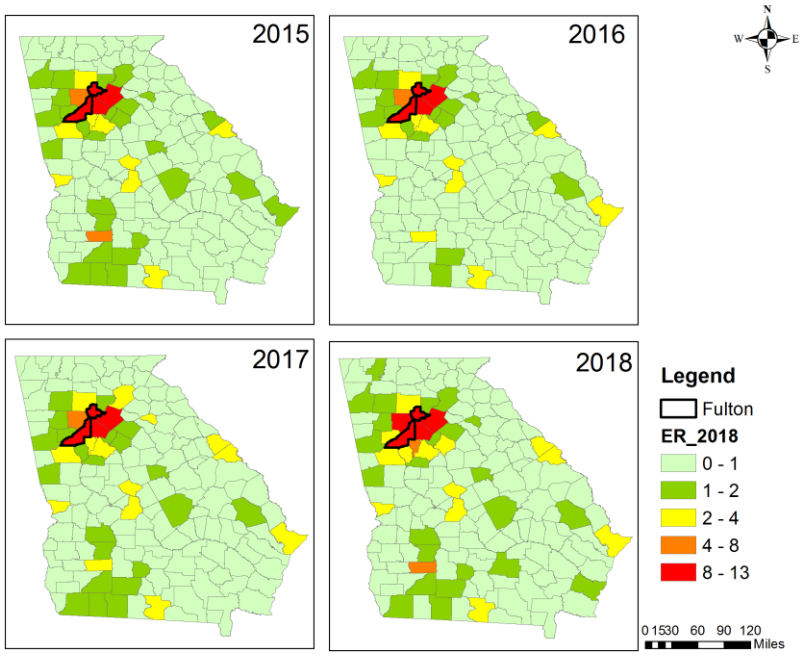

Figure 6-11 Asthma-related ER visits by the county in Georgia from 2015 to 2018, first four months

The parameters in health impact function include uncertainty that should be considered. We used the β from a wildfire epidemiological studies. It may be different compared to a prescribed fire specific case. An epi-study focuses on Atlanta area (Krall et al. 2017) shows a positive association (Relative Risk = 1.006) between respiratory disease and biomass burning source. Considering the sources of $PM_{2.5}$ in Georgia, the biomass burning is mainly prescribed burning emissions. Besides the uncertainty in coefficient β, there are other uncertainties in estimating prescribed burning related health impact. Previous studies show that the measured $PM_{2.5}$ emissions from prescribed burning is about 90% of modeled emission (Davis et al. 2015). In our case, BlueSky estimated $PM_{2.5}$ emissions is about 60% of measured emissions. The horizontal and vertical allocation of fire emissions in model could lead to about 20% uncertainty in the final pollutant concentration estimation (Garcia-Menendez et al. 2014). Inaccurate simulated wind speed

and wind direction may also bring uncertainty as high as 100% (Garcia-Menendez et al. 2013). Fire duration is also hard to be determined due to lacking of post-fire information.

6.4 Conclusions

Application of the data fusion method to adjust fire impacts improve exposure fields for health analysis. However, the lack of observations sites can lead to missing the major fire impacts on air quality, though this is captured in the simulation. Fusing the two can lead to overly reduced fire impact estimates, though the smoke plume is still captured. The lack of monitoring sites can be alleviated, in part, by using inexpensive sensors. Using adjusted fire impact which comes from the data fused exposure fields that multiplies the ratio between fire impact and total $PM_{2.5}$ from CMAQ-DDM to compare with observations could help researchers target the day and area that prescribed fire has a major impact on the local air quality even if the observation is low and also distinguish those days from the other days when fire impact is low. Those days and areas should be investigated further in epidemiological studies to find the relationship between health effects and prescribed burning.

Here we used the β from a wildfire epidemiological studies. Lacking epidemiological studies to provide prescribed burning-specific concentration-response functions ($\beta's$) is a weakness of this health analysis. However, according to the health impact function, changing of β would only change the level of health impact and the exact number of prescribed burning impacted people, the spatial and temporal variations of prescribed burning impact on public health would not change. The variations of prescribed burning activity lead to the monthly and locations differences on health impact. Those

results illustrate that it is important to distinguish seasons and areas when studying prescribed burning and its health impacts. While southern Georgia has the higher density of prescribed burns, the greatest health impacts, in terms of absolute number of asthma-related ER visits, are found in the southwestern, central, and east-central Georgia, as well as the Atlanta MSA, given the much larger populations along with moderate or high levels of prescribed burning impacts. Atlanta MSA is the most fire impacted area with the largest number of ER visits due to the large population there. Although southwestern, central, and east-central Georgia have larger health impacts, the number of ER visits related to asthma is smaller compared to Atlanta MSA.

Prescribed burning and wildfire impacts will become an increasing fraction of $PM_{2.5}$ exposures in the future as controls continue to reduce emission from other sources while emissions from these two sources are expected to increase. Here we found an increasing trend in health impact due to prescribed burning. Strategic use of prescribed burning, however, can be used to reduce human exposures by conducting prescribed burnings on days leading to lower exposures and also reduce wildfire-related exposures. Not only Georgia but the entire southeastern U.S. that houses some of the most vulnerable communities in the nation is more likely to experience high and frequent smoke exposure in comparison to the other parts of the country due to increasing prescribed burning emissions. A better understanding of the contributions of prescribed burning to human health is important, especially to the people who are affected by prescribed burning directly.

6.5 Acknowledgment

This publication was made possible in part by funding from the Joint Fire Science Program under grant number 16-1-08-1 and NASA Applied Sciences Program under grant number NNX16AQ29G. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the supporting agencies. Further, the US Government does not endorse the purchase of any commercial products or services mentioned in the publication. We thank everyone who helped us with the data, in particular:

- Mr. Michel Klein, Drs Stefanie Sarnat and Paige Tolbert of Emory University for providing ER visits data.

CHAPTER 7. CONCLUSION AND FUTURE WORK

7.1 Conclusions

The findings in this dissertation help us better understand the impact of prescribed burning and its impact on air quality and human health. The findings are summarized below.

We found the limitations of current satellite-derived products in estimating burned areas of small fires like prescribed burning by comparing them with permit record datasets. Satellite-derived products have coarse resolution relative to the size of prescribed burning. They can capture the spatial variation of the prescribed burning activity at the county level, but not the level of burns. Clustered fires are easier to detect than an isolated one. However, satellite products may misinterpret those small fires as a combined large fire. Those products need to improve the inaccuracy in detecting small fires by incorporating new technologies, like newly launched satellites with finer spatial and temporal resolution and algorithms used in calculating the burned area. Due to the uncertainty in satellite-derived products, emission inventories and other research that use the burned area data as input should adjust them accordingly if they focus on small fires.

A novel source apportionment method (DASI) is developed and applied to split combined prescribed burning impact obtained from the model simulation. The application of the method to get the single fire impact could help land and air quality managers quickly decide whether the burn permits should be restricted or if more permits could be issued. The method could also be applied to source apportionment of other emission sectors by

splitting the pollutant concentrations generated by the chemical transport model. Considering the uncertainties of current source apportionment methods based on mass balance, which do not take the chemical reactions into account, and the computational time and space the chemical transport models may use, our method provides a more efficient way to conduct source apportionment with less uncertainty compared to receptor modeling methods. However, more applications are needed to compare the results with other source apportionment methods to evaluate the new method.

The feasibility of a low-cost PM sensor to detect fire impact has also been evaluated. Low-cost sensors have the advantage of being lower cost, portable, and easy-to-use. They can be used as a supplemental tool to measure ambient concentration in areas lacking regulatory monitors. Four low-cost PM sensors have been deployed in southwestern Georgia, one of the most active prescribed burned area in the southeastern U.S., from May 2017 to June 2018. We found that the low-cost PM sensor we utilized could be used to detect fire impact. They can provide back-up measurements when regulatory monitors stop working. They can also capture the fire impact missed by the single monitoring site nearby, according to model simulation analysis. However, due to the highly non-homogeneous distribution of $PM_{2.5}$ concentrations from fire impact, spatial variations cannot be captured even with a 4-km resolution simulation. Uncertainties in wind speed due to a systematic bias at nighttime in the WRF model and wind direction will also limit the accuracy of the simulations. Therefore, in order to generate a more accurate estimation of pollutant exposure fields to prescribed burning, we recommend fusing observations from a dense network of low-cost sensors and model simulations using the data fusion method.

The data fusion method has been applied from 2006 to 2008 over North Carolina to develop the spatiotemporal fields of $PM_{2.5}$ and its species and gaseous pollutant concentration fields as well. The method fuses observations from monitor sites and Community Multiscale Air Quality (CMAQ) model pollutants fields to generate air pollutant exposure fields for health studies. The optimized fields resemble the CMAQ-observation fields near the monitor sites and the scaled CMAQ fields far from the monitor sites. The final outputs are consistent with the observation variations and capture spatial information by the air quality model. Inter-comparison of $PM_{2.5}$ exposure fields using data fusion and two other methods including satellite-derived AOD data show that the data fusion method provides the strongest correlation and lowest errors in this study. However, as each method has its own pros and cons, we should choose them according to different situations as appropriate.

The data fusion method has also been used to generate exposure fields from prescribed burning impact for the first four months from 2015 to 2018 in Georgia for quantifying the health impact due to the fires. We provided a way to choose the days and areas that prescribed burning has a major impact on the local air quality which is important for epidemiological studies to explore the relationship between prescribed burning and health effects. The health impacts results show a strong spatial and temporal variation of prescribed burning. However, lacking epidemiological studies for prescribed burn smoke to provide the coefficient β that coming from the concentration-response function may weaken our health analysis. Further health studies are needed to investigate the impact of prescribed burning on human health.

7.2 Future work

This dissertation has provided a basic understanding of how prescribed burning affects air quality and human health by using measurements and model simulations and could be treated as a foundation for additional studies in the areas of prescribed burning and source apportionment.

Unlike Georgia and Florida, which have relatively complete prescribed burning permit records in their system, there are some active prescribed burning states that do not have this kind of system. The permit records themselves also have uncertainty. It would be useful to be able to use satellite-derived estimates more widely, particularly in areas that do not have systems that are as extensive or as available as Florida and Georgia. Statistical and machine learning methods (e.g. logistic regression model, Bayesian method, random forest) could be used to calibrate the current satellite-derived product. Taking Georgia as the test state which has longer (2013 to 2018) permit records, we can use the first five years (2013 to 2017) to train the model to level up the burned area from satellite-derived products. For the areas that have permit burns but not detected by satellite, we can use methods mentioned previously to calculate the possibility that whether it burns or not. For instance, for the grid that burns every year according to permit records, even the satellite does not capture the burn there, we can treat that grid as 100% burn grid. For the grid that burns with a cycle (e.g. 2 years period), we can estimate the possibility that it burns in the current year. Finally, we can treat the area that has the burn possibility (e.g. 80%) as the real burned area. The calibrated results will be useful for emission estimation and improving model simulations on quantifying the impact of prescribed burning. Also, evaluation of burned area using newly launched satellites such as GOES-16 could be conducted. Hazard Mapping System now incorporates multiple satellites to generate a

burned area product. We can choose a recent year (2018) to do repeat comparison between permit records and satellite-derived product to find whether there are improvements using newly launched satellites.

The new source apportionment method could be further applied to the forecast system. With information on the marginal contribution of each burn to the air quality downwind, the following dynamic management protocol can be employed (Figure 7-1). The burns that have minimal impacts are permissible, while the ones that contribute a lot may have to be denied. In this case, burn A contributes almost nothing to the regional peak $PM_{2.5}$ concentration. Burn B contributes a little bit more, but still not very much. Burn C contributes still more, and so on up through burn H. If all the burns A-H were permitted, then the region would exceed the NAAQS. If only burns A-F were permitted, however, $PM_{2.5}$ concentrations would stop short of the NAAQS. This approach might maximize the amount of land that could be burned without going over the NAAQS. Burning any more (i.e., burns G and H) would put the region over the NAAQS.

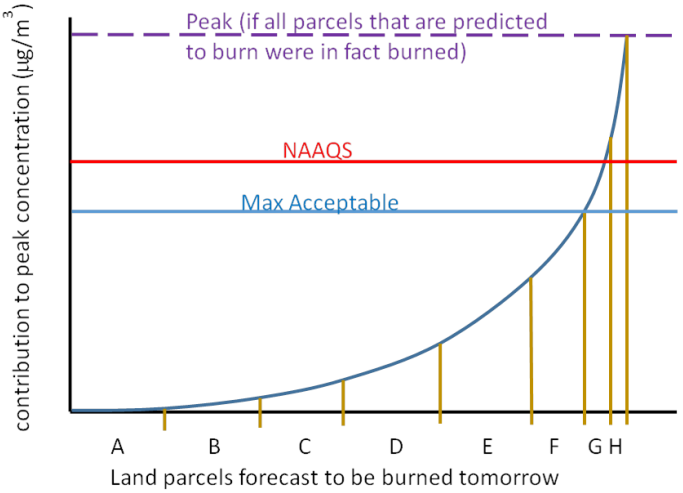

Figure 7-1 Illustration of how forecasts of the air quality impacts of individual burns can be employed for dynamic air quality management (Odman et al. 2017)

DASI could also be evaluated by comparing with other source apportionment methods (CMB, CMAQ-DDM). The chemical transport model and dispersion model could be changed accordingly.

A dense network including other types of low-cost sensors can be built in southwestern Georgia or other places to further investigate the performance of those sensors in detecting fire impact or other sources. The on-road mobile emissions sector is another important source of air pollution. It would be quite interesting to look at the near-road air pollution concentrations using low-cost sensors. The spatial variation provided by the near road low-cost sensors network could benefit the health studies to better estimate the exposure fields and minimize the uncertainty in CTM using data fusion method.

There are many new methods developed and improved to provide the spatiotemporal exposure fields for health studies such as satellite-retrieved AOD included methods, improved Land Use Regression model, hybrid methods that incorporate AOD data, monitor observations, land use variables and CTM simulations in the last decade. All those methods have their pros and cons and should be evaluated by inter-comparison for the same domain (CONUS) at the same spatial resolution. Those methods should also be assessed in applications to health studies to see how different health results may respond to different methods, especially at different resolutions and to answer the question of how fine the exposure fields should be for different type of health studies considering the time and effort spent, to find the balance between appropriate resolution of exposure fields and reasonable health results.

In sum many, quantifying the impacts of prescribed burning from each fire, especially one large fire or a cluster of small fires at the same time, on air quality and human health in the Southeast is important. Prescribed burning is a preferred land management tool but also a prominent source of air pollution in the U.S., and exposure to fire smoke is a growing health concern.

APPENDIX A. CHAPTER 2 SUPPLEMENTAL INFORMATION

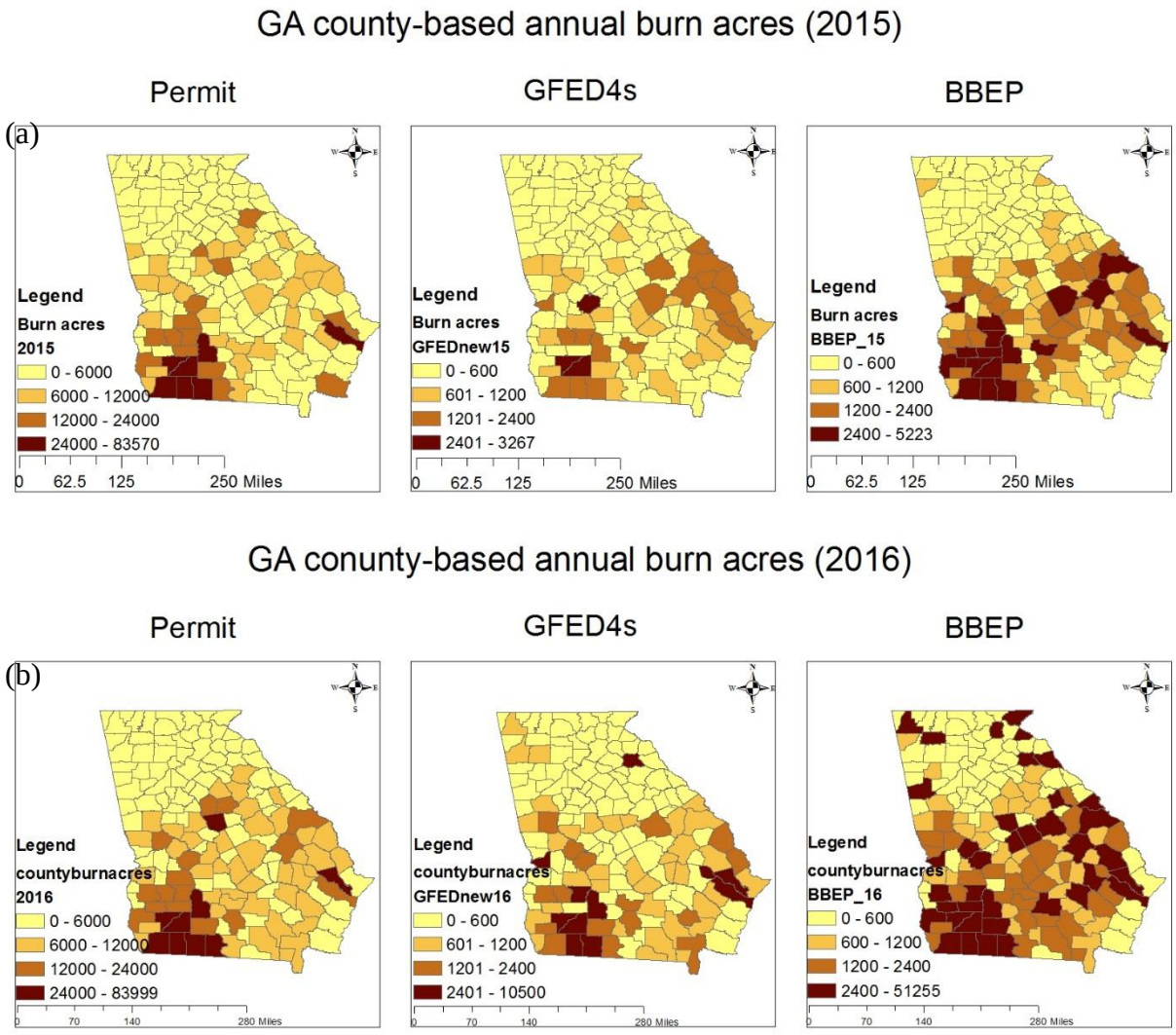

Figure A - 1 Total burned areas of first four months of the year by county in Georgia: a) 2015, b) 2016. The scale for permit records is 10 times larger than the scale for GFED4s and BBEP.

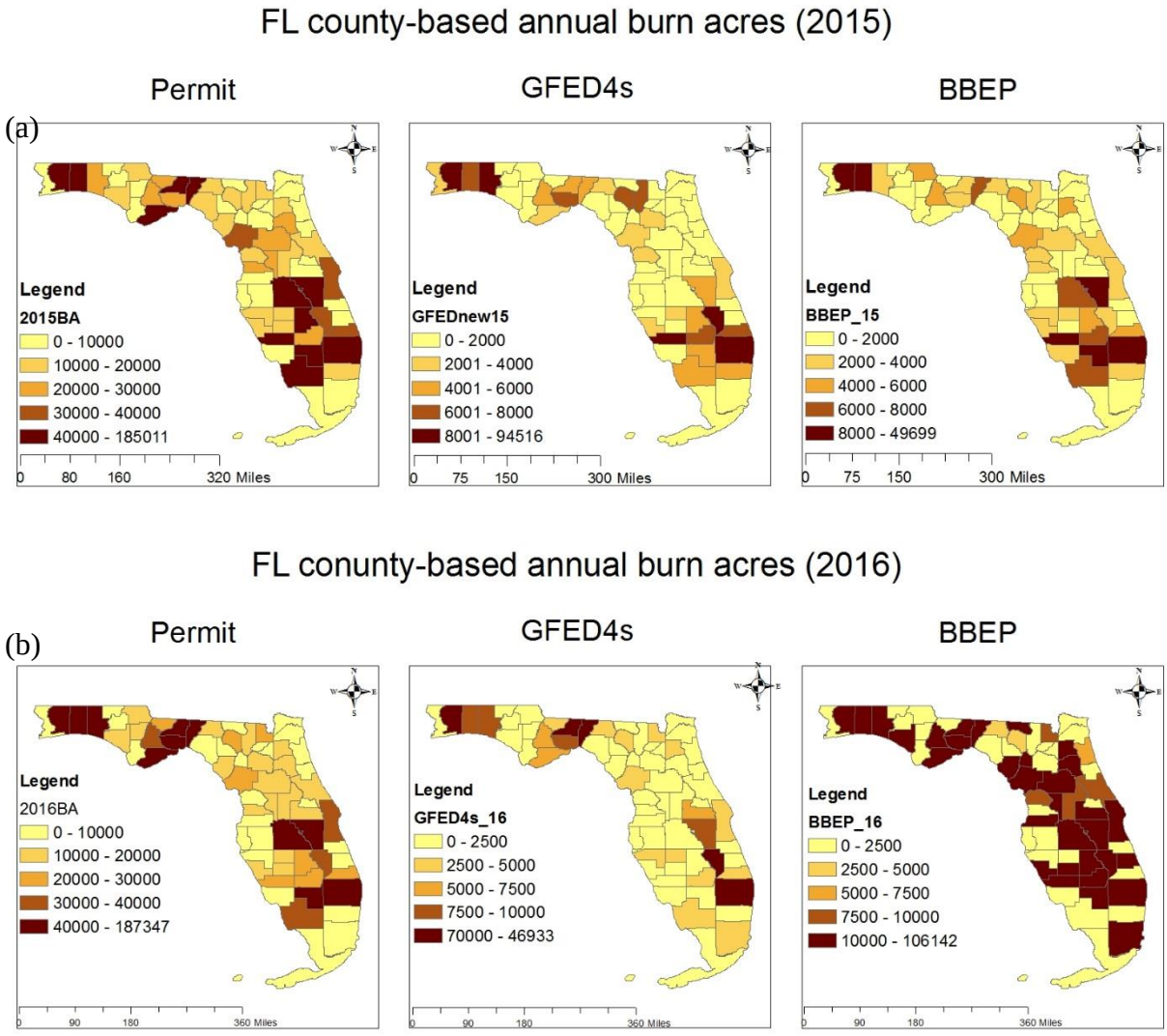

Figure A - 2Total burned areas of first four months of the year by county in Florida: a) 2015, b) 2016. The scale for permit records is 4 or 5 times larger than the scale for GFED4s and BBEP.

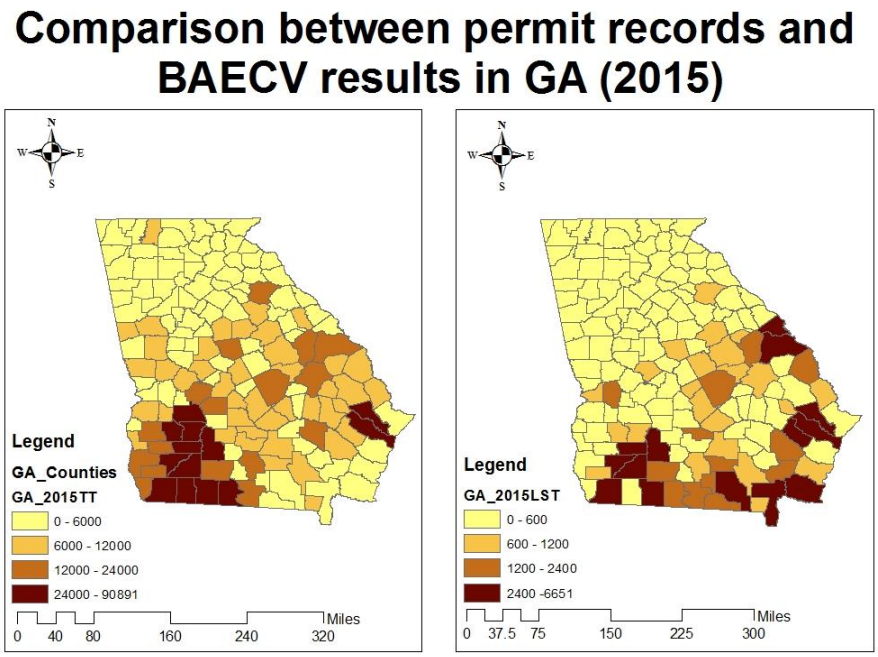

Figure A - 3 Comparison between permit records (left) and BAECV results (right) in Georgia (2015). The scale for permit records is 10 times larger than the scale for BAECV.

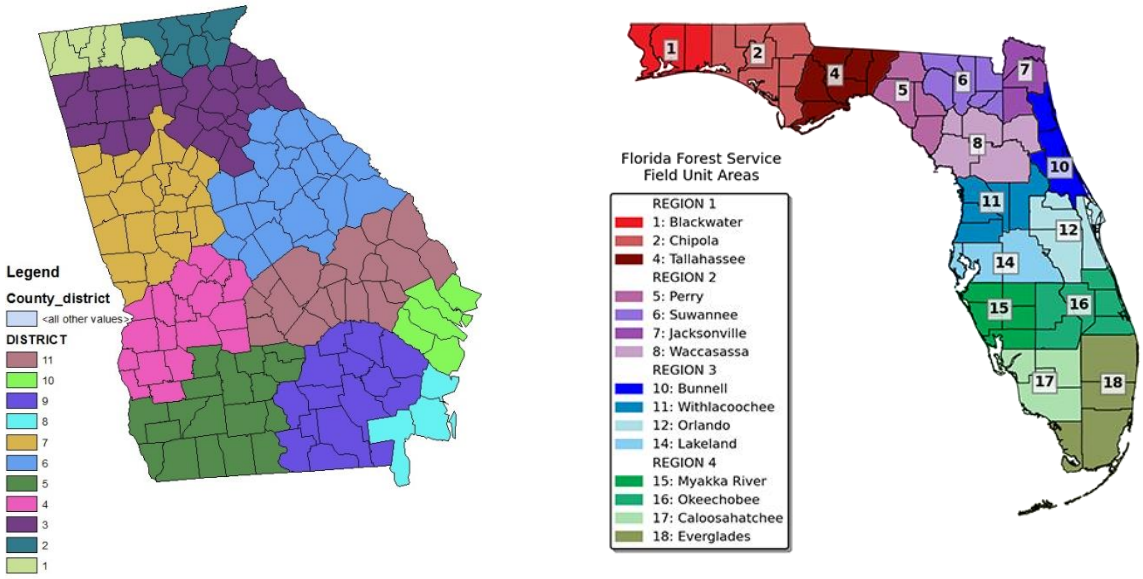

Figure A - 4 Fire Districts in Georgia (a) and Florida (b)

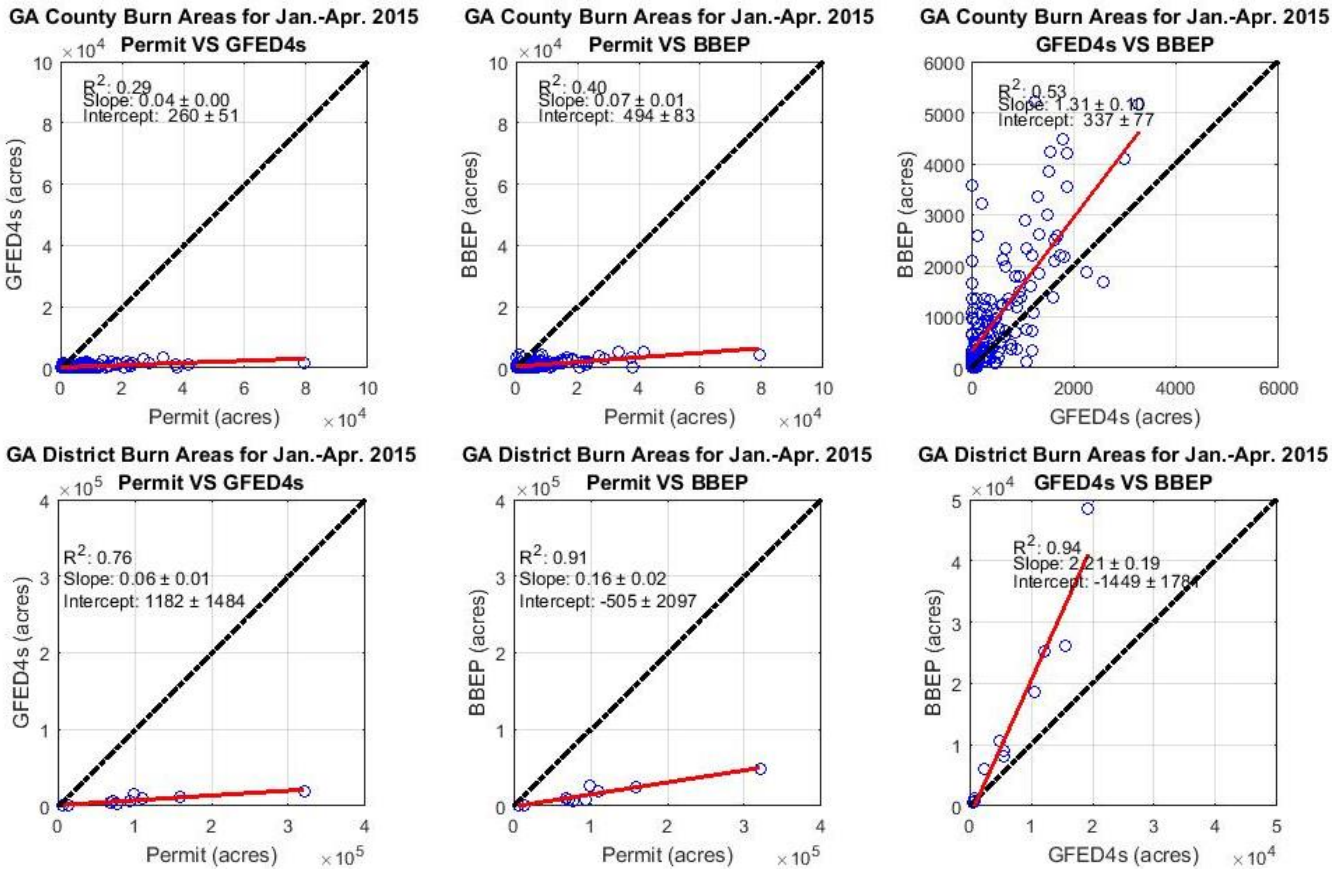

Figure A - 5 Inter-comparisons of permit record, BBEP and GFED4s burned areas in Georgia for the first 4 months of 2015: County totals (top row) and district totals (bottom row).

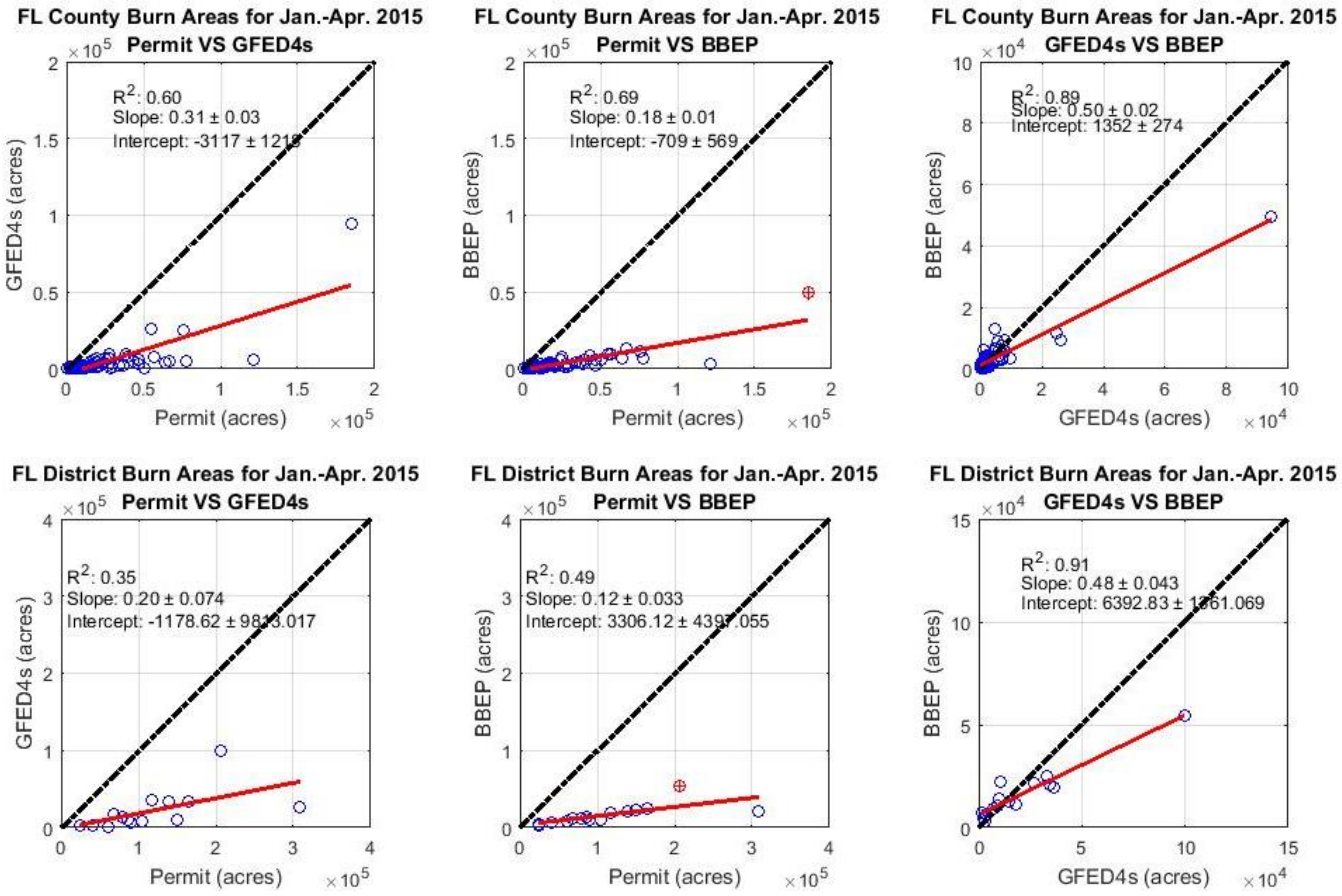

Figure A - 6 Inter-comparisons of permit record, BBEP and GFED4s burned areas in Florida for the first 4 months of 2015: County totals (top row) and district totals (bottom row). The red dots in the middle panels represent Palm Beach County and District 18 where sugarcane burning dominates.

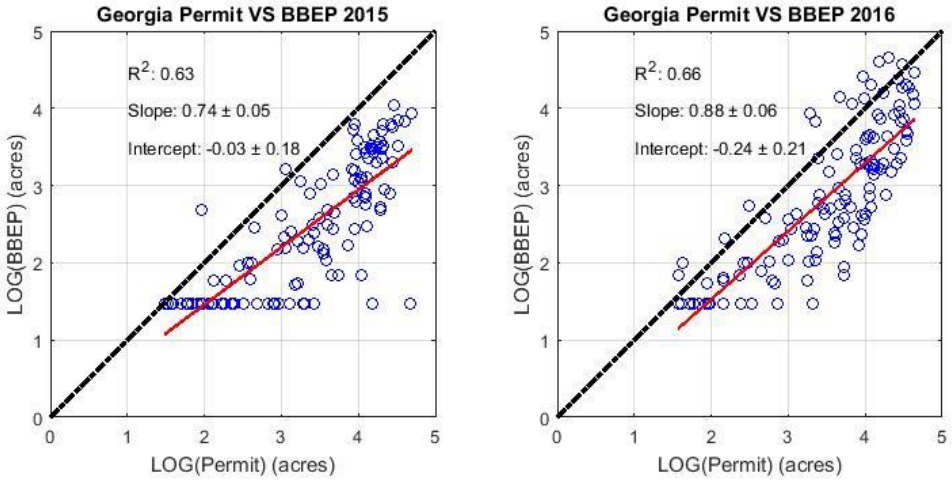

Figure A - 7 Comparison of daily state total burned areas for the first four months of 2015 and 2016 in Georgia: BBEP versus permit record data (Log-transformation).

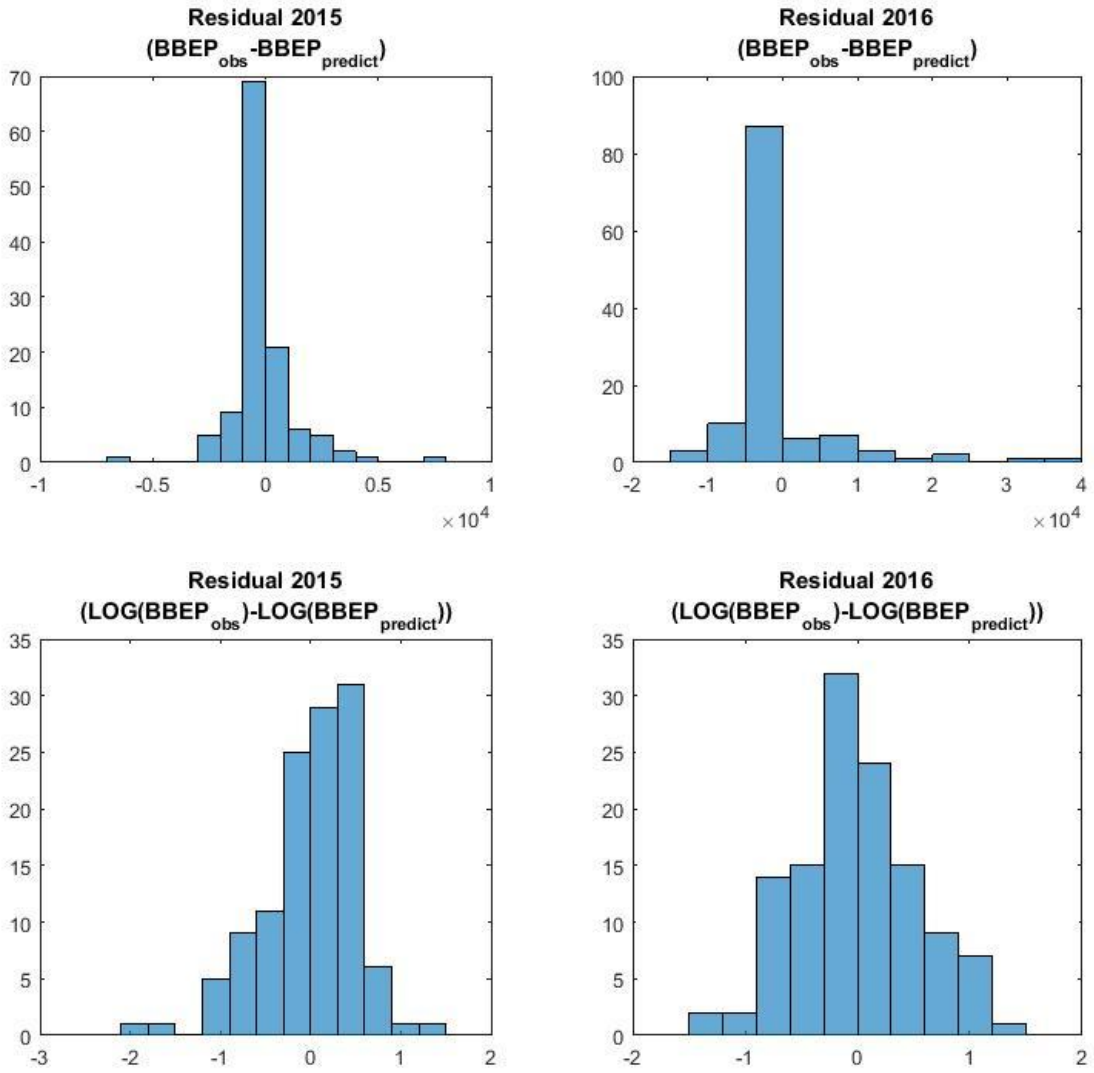

Figure A - 8 Residuals of permit records and BBEP (top) and log-transformation permit records and BBEP (bottom) in GA.

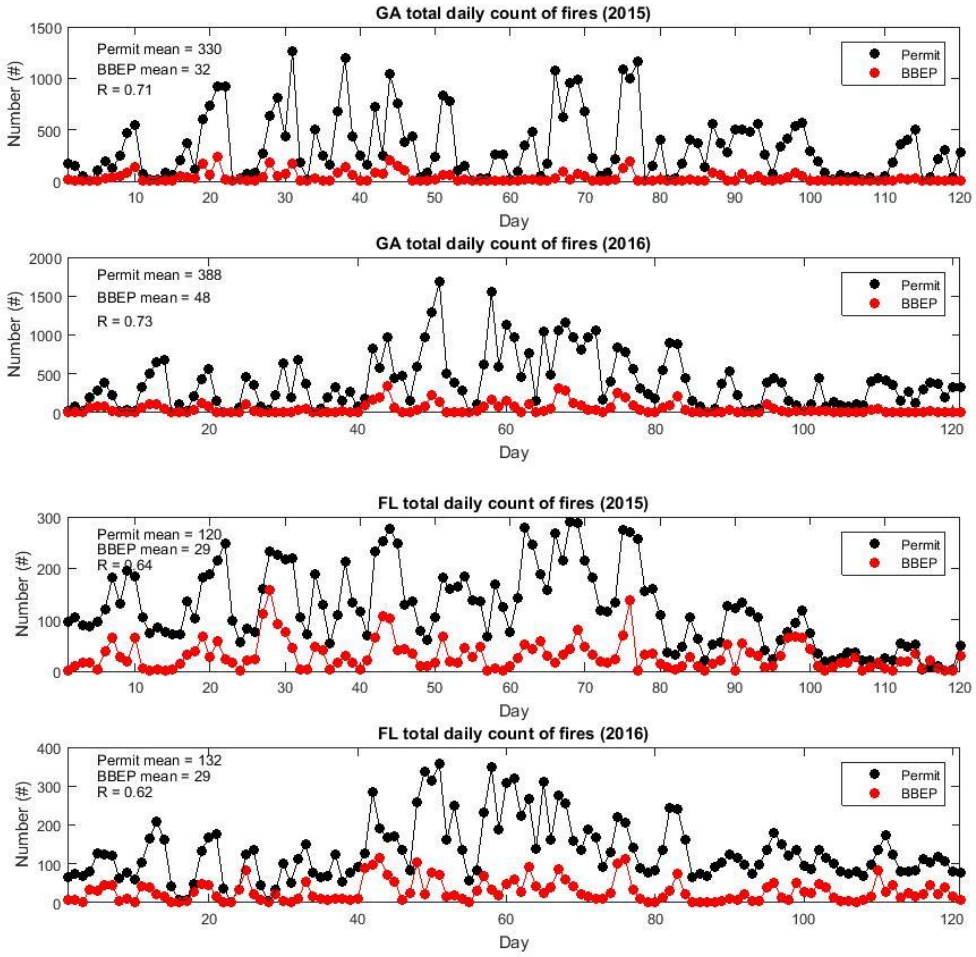

Figure A - 9 Comparison of daily count of fires between permit record data and BBEP in Georgia (top) and Florida (bottom) for the first 4 months of 2015 and 2016.

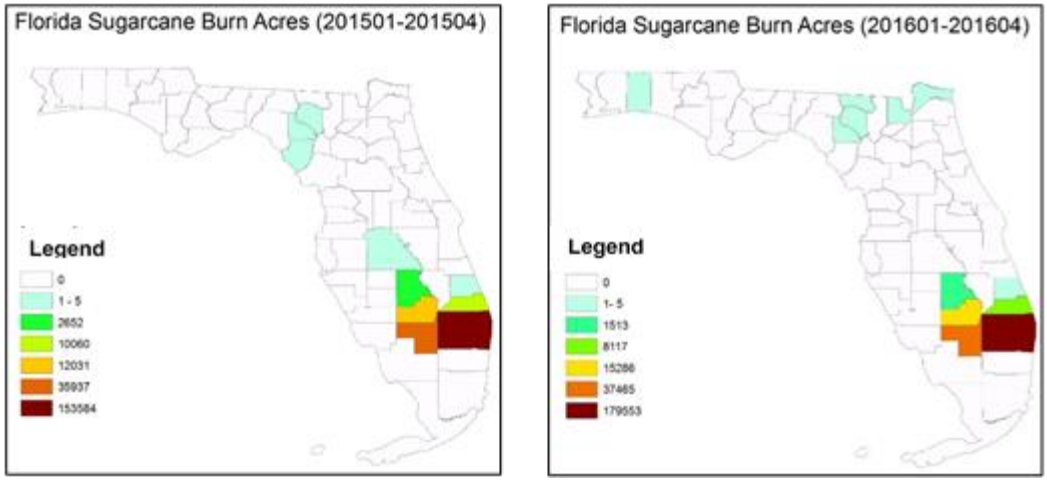

Figure A - 10 Total burned areas of permitted sugarcane burns by county in Florida during the first four months of 2015 (left) and 2016 (right).

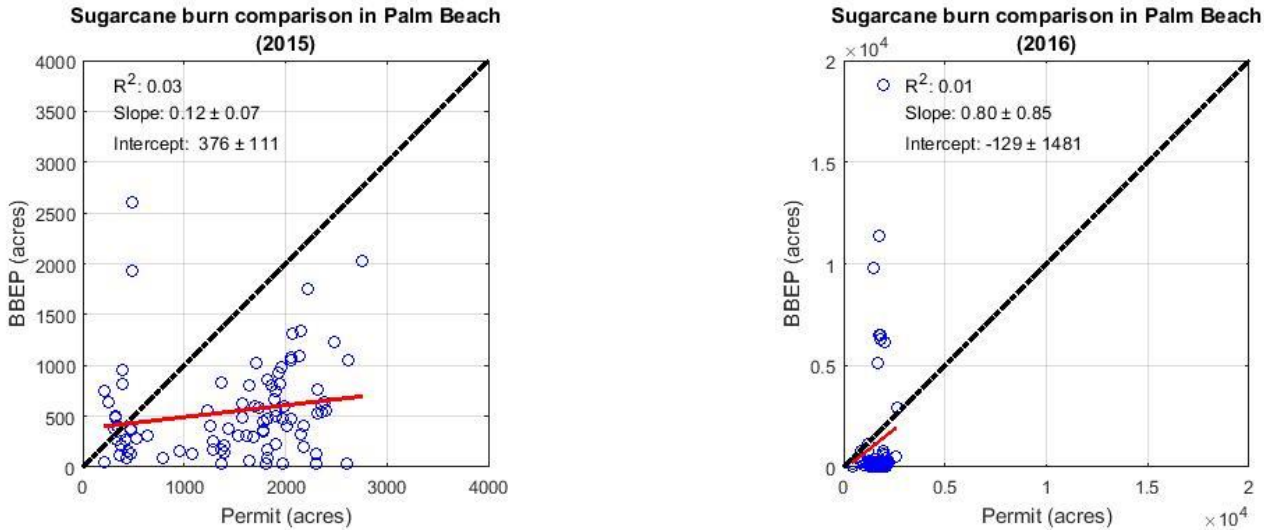

Figure A - 11 Comparison of burned areas between permitted sugarcane burns and BBEP-detected fires in Palm Beach County, Florida for the first four months of 2015 (left) and 2016 (right).

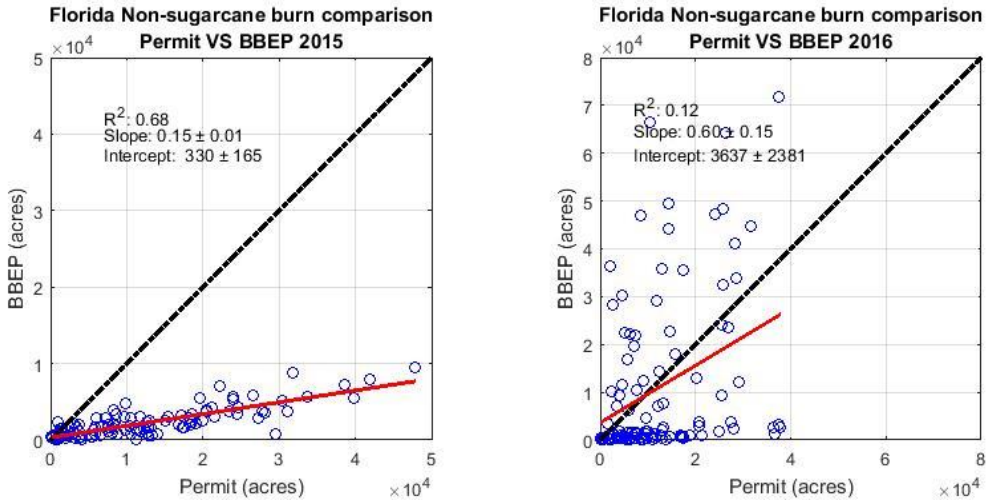

Figure A - 12 Comparison of daily state total burned areas between permitted non-sugarcane burns (i.e., all burn types except sugarcane) and BBEP-detected fires in Florida for the first four months of 2015 (left) and 2016 (right)..

Table A - 1 Size distribution of fires according to burn permit records and BBEP in Georgia for 2015 and 2016

Georgia	Permit		BBEP	
Year / Size	2015	2016	2015	2016
0-5	47.0%	46.3%	0.0%	-
5-10	9.6%	9.4%	1.5%	4.8%
10-25	16.5%	16.5%	14.3%	13.4%
25-50	10.7%	11.0%	76.6%	78.3%
50-100	11.3%	12.1%	2.2%	0.3%
100-250	3.9%	3.9%	5.0%	2.1%
250-500	0.7%	0.7%	0.3%	0.2%
500-100	0.2%	0.1%	0.0%	-
1000+	0.1%	0.1%	-	0.8%
Total count of fires	39602	46958	3872	5814

Table A - 2 Size distribution of fires according to burn permit records and BBEP in Florida for 2015 and 2016

Florida	Permit		BBEP	
Year / Size	2015	2016	2015	2016
0-5	17.0%	19.5%	-	-
5-10	4.1%	3.6%	1.7%	2.8%
10-25	12.9%	12.6%	15.9%	15.8%
25-50	23.9%	25.7%	48.0%	67.5%
50-100	21.3%	20.3%	10.4%	2.4%
100-250	11.7%	11.1%	21.3%	5.7%
250-500	4.3%	3.8%	2.4%	0.8%
500-100	2.9%	1.9%	0.3%	0.2%
1000+	1.9%	1.5%	0.1%	4.7%
Total count of fires	14389	15933	3526	3454

APPENDIX B. CHAPTER 3 SUPPLEMENTAL INFORMATION

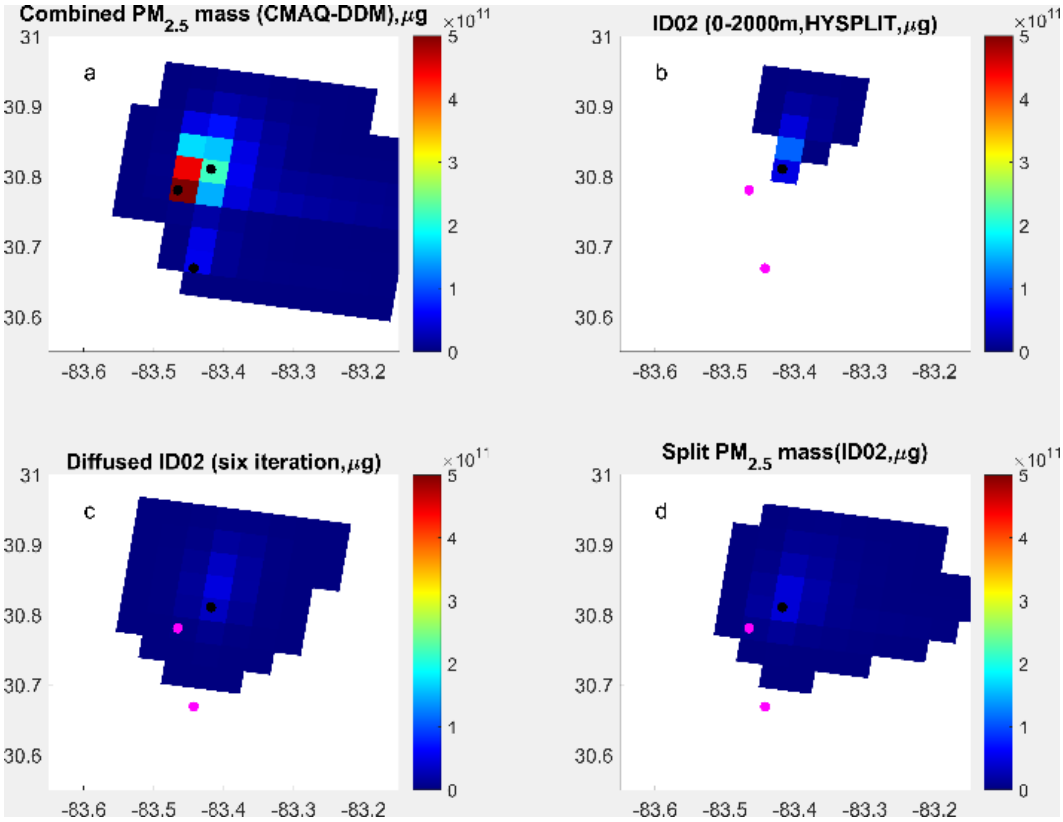

Figure B - 1 Split fire impacts based on diffused HYSPLIT fields of fire ID02

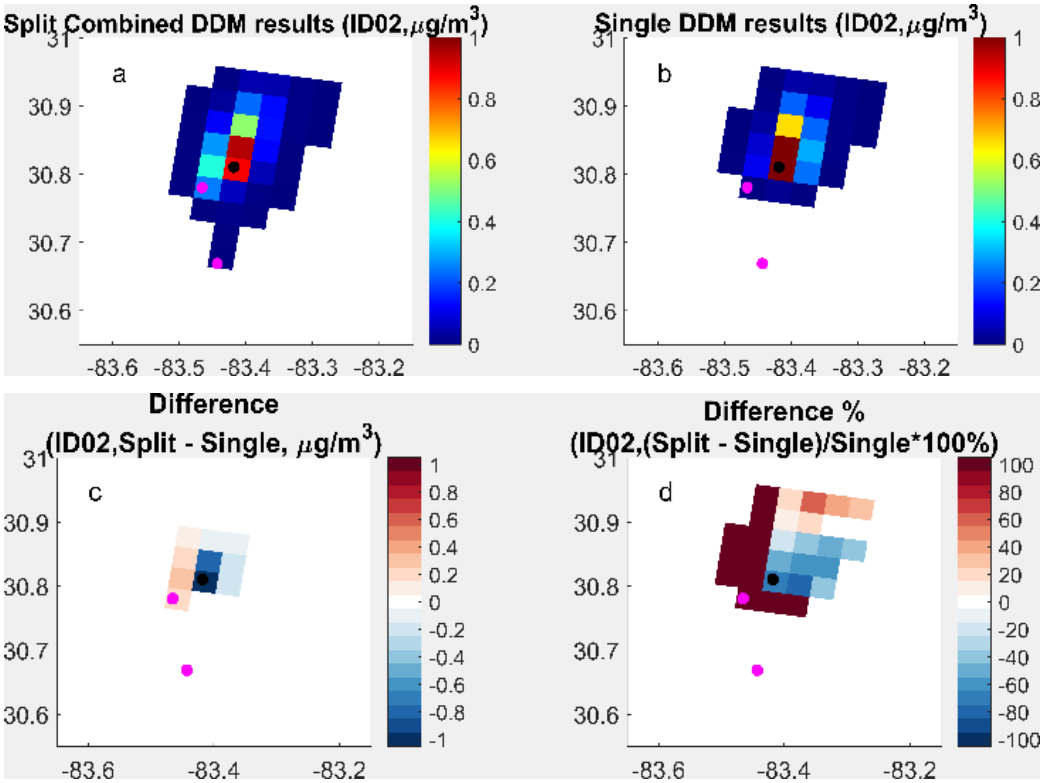

Figure B - 2 Comparison between split single fire impact and single fire impact from CMAQ-DDM of fire ID02 at ground layer

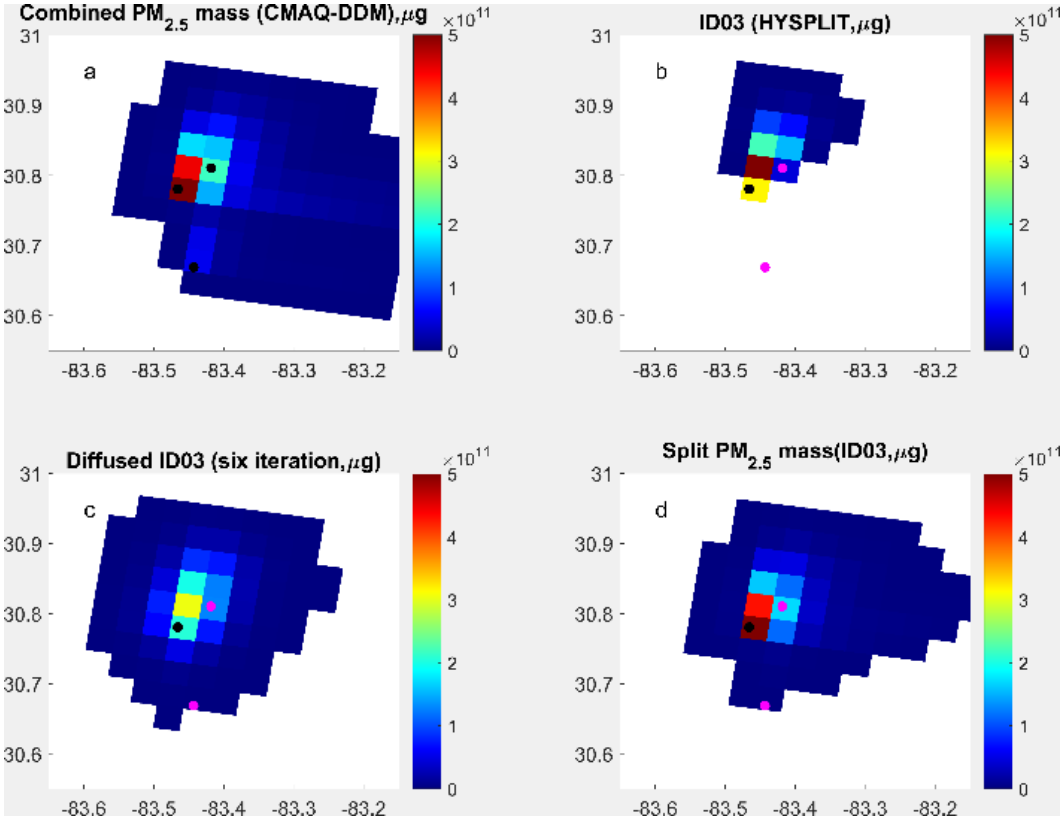

Figure B - 3 Split fire impacts based on diffused HYSPLIT fields of fire ID03

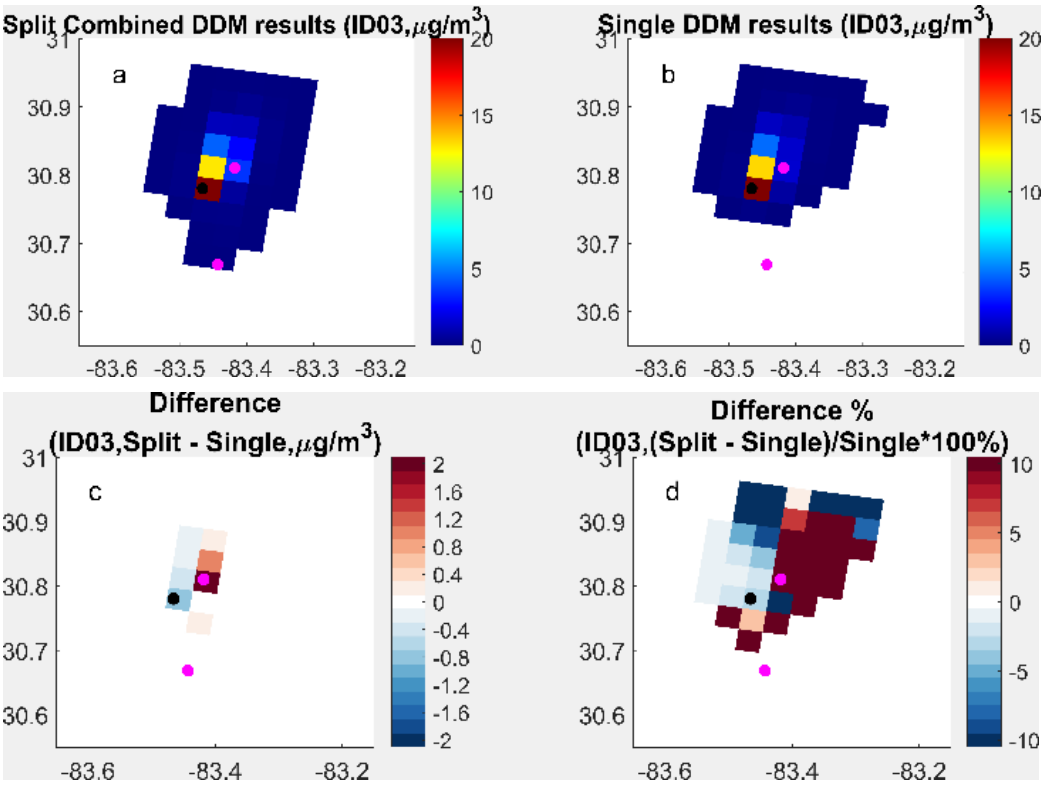

Figure B - 4 Comparison between split single fire impact and single fire impact from CMAQ-DDM of fire ID03 at ground layer

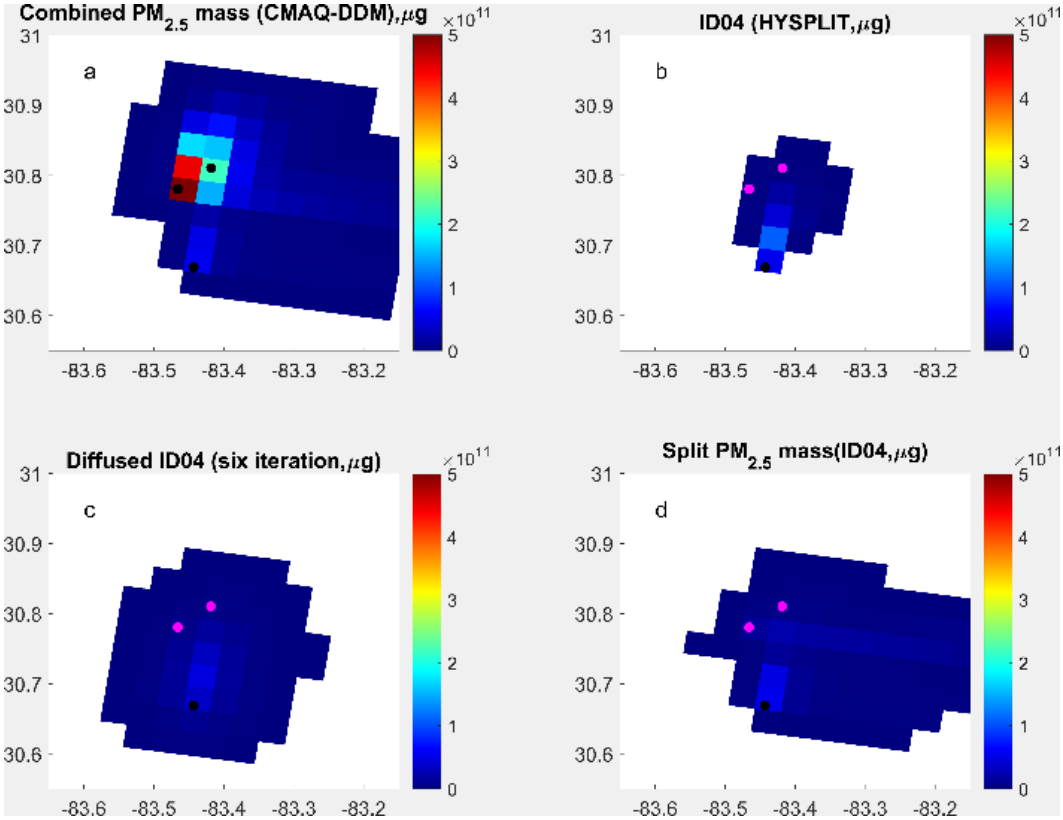

Figure B - 5 Split fire impacts based on diffused HYSPLIT fields of fire ID04

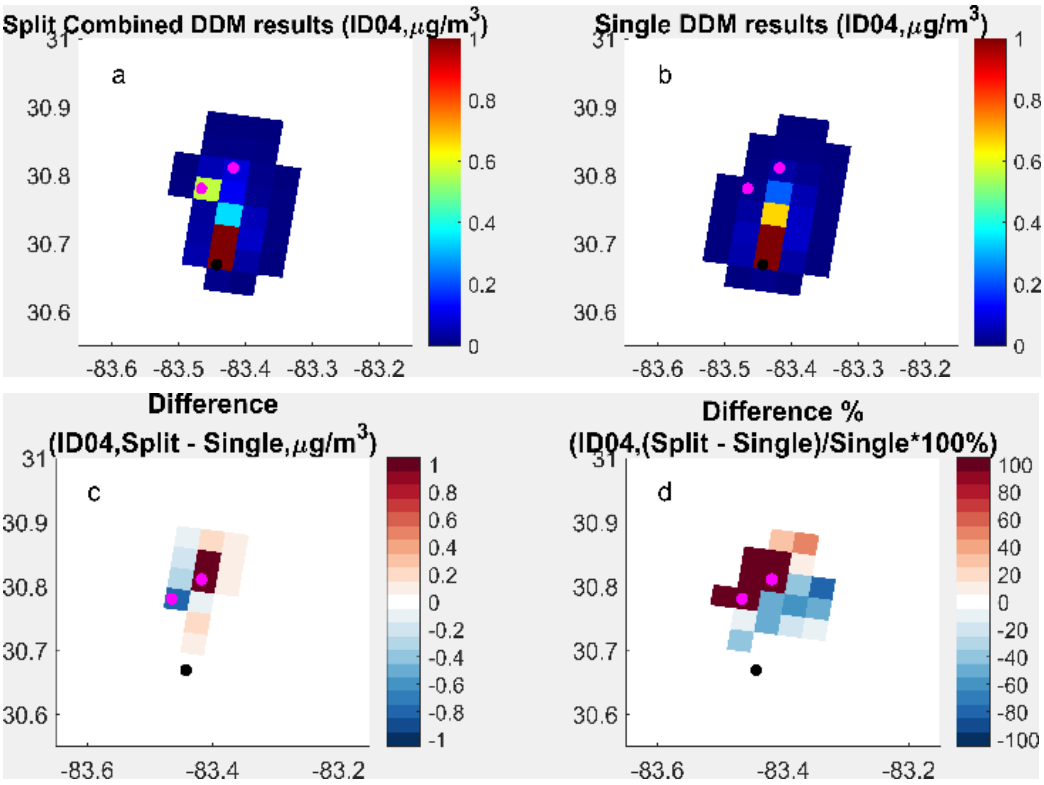

Figure B - 6 Comparison between split single fire impact and single fire impact from CMAQ-DDM of fire ID04 at ground layer

APPENDIX C. CHAPTER 4 SUPPLEMENTAL INFORMATION

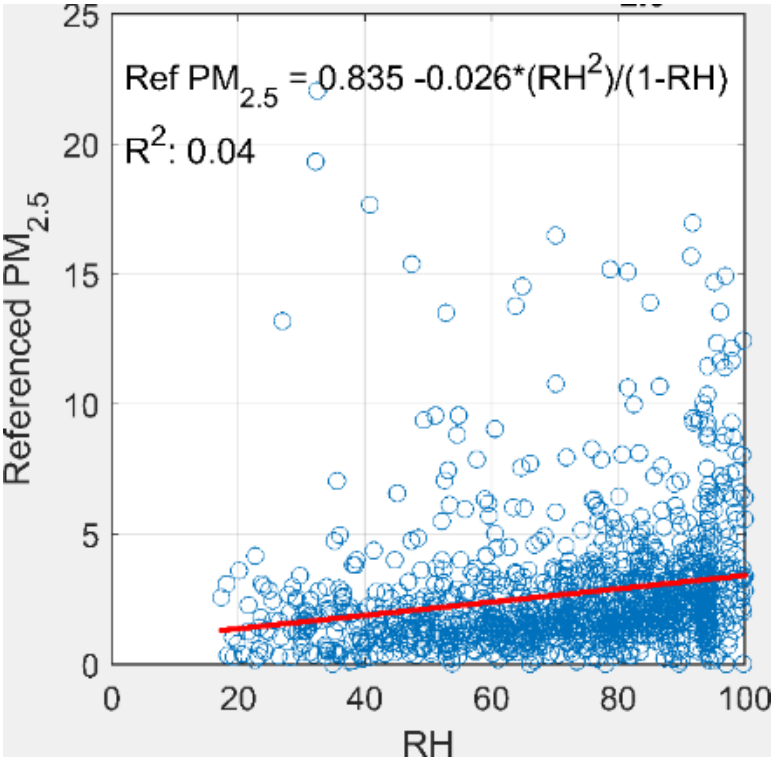

Figure C - 1 Correlation between relative humidity from Southwest Georgia Regional Airport Site, and $PM_{2.5}$ from the BAM at the Albany EPD site

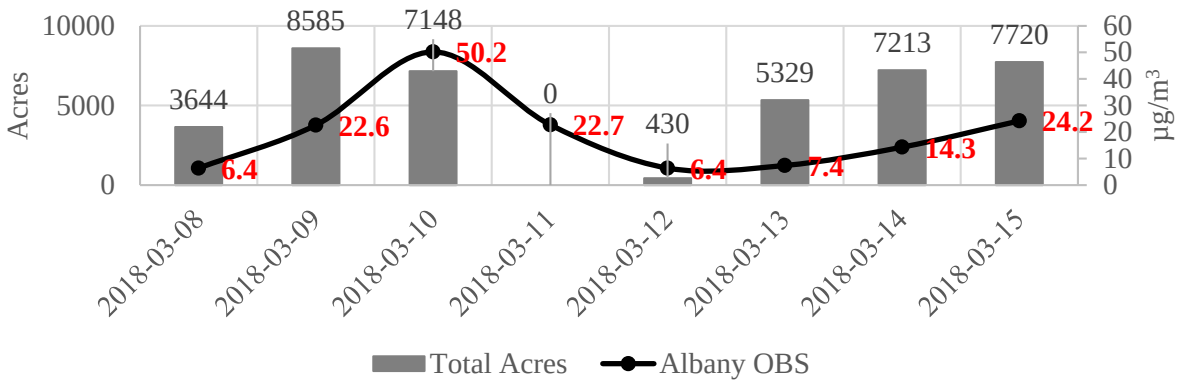

Figure C - 2 Daily total burn acres from GFC permit data at low-cost sensor deployed and surrounding counties and observation from EPD site (red text: daily $PM_{2.5}$ concentration)

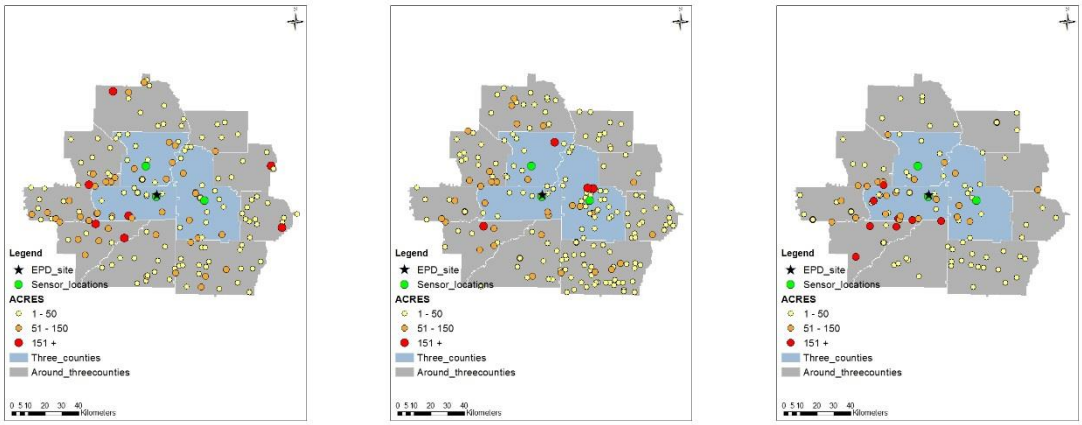

Figure C - 3 Permitted burns on March 9 (left), March 10 (middle) and March 13, 2018 (right): red circles represent fires larger than 150 acres

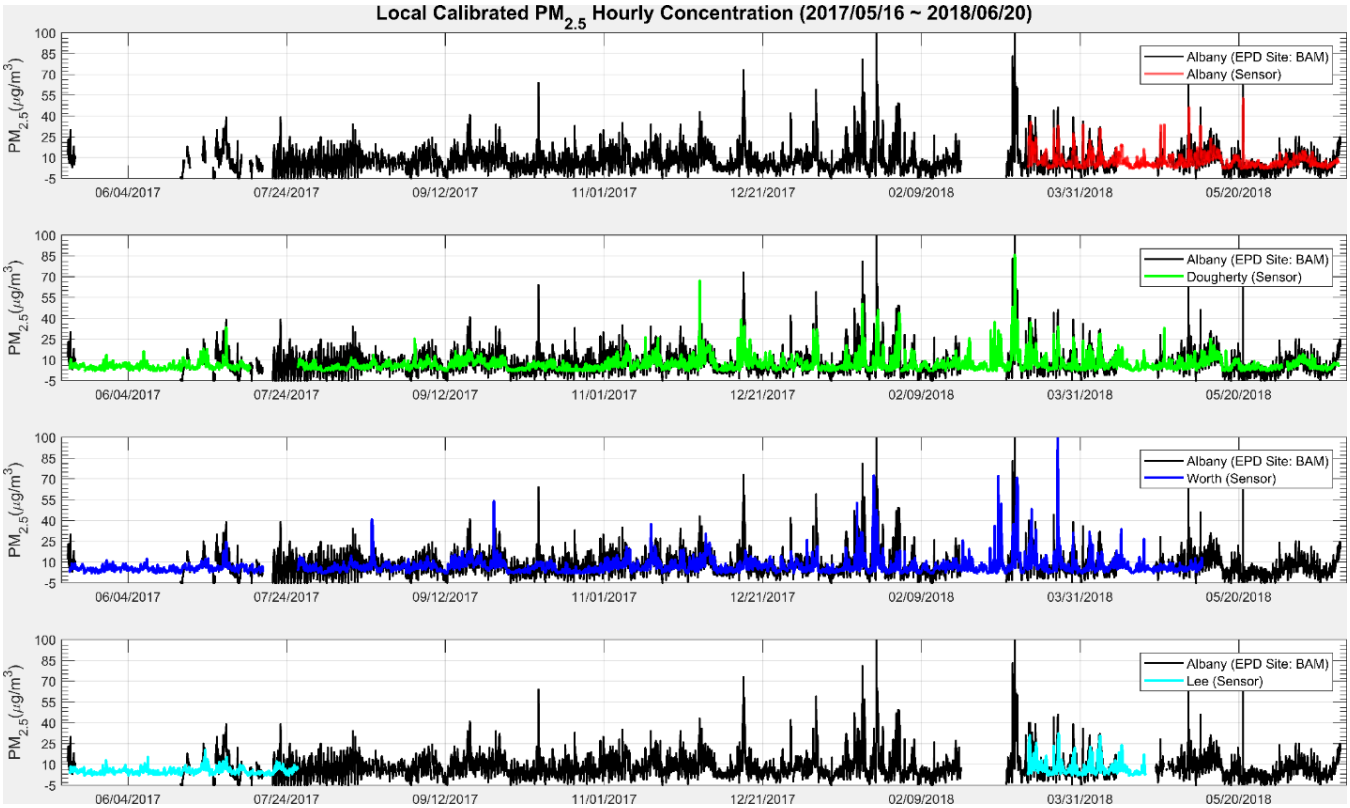

Figure C - 4 Hourly $PM_{2.5}$ concentrations from low-cost sensors and BAM at GA EPD site (Albany) from May 16, 2017 to June 20, 2018

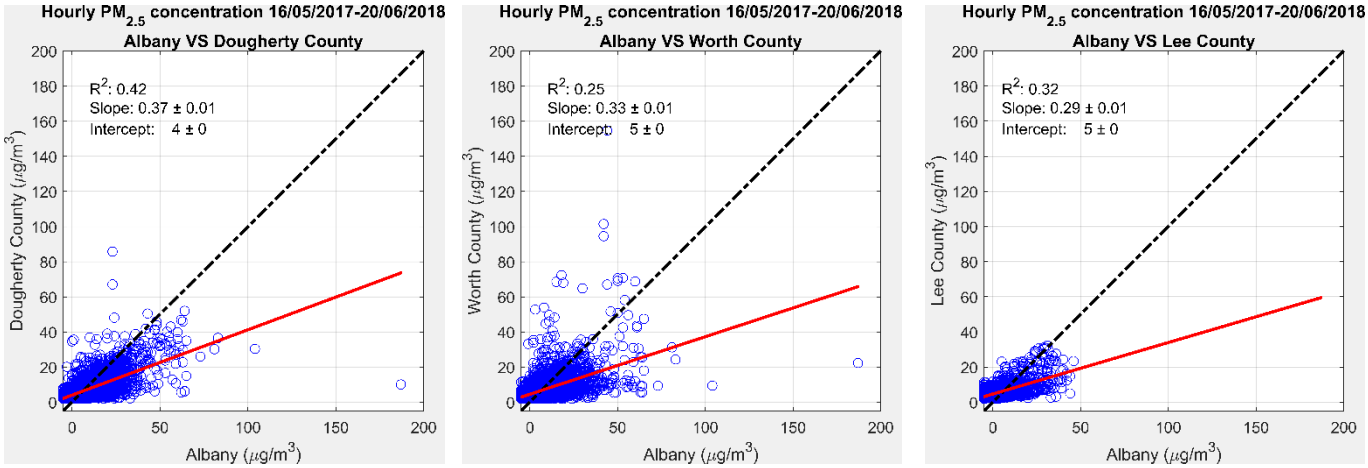

Figure C - 5 Comparison of hourly $PM_{2.5}$ concentrations between high school low-cost sensors and BAM at GA EPD site (Albany) from May 16, 2017 to June 20, 2018 (Lee County sensor was down from July 28, 2017 to March 13, 2018 and April 21, 2018 to June 20, 2018; Worth County Sensor was down from May 09, 2018 to June 20, 2018)

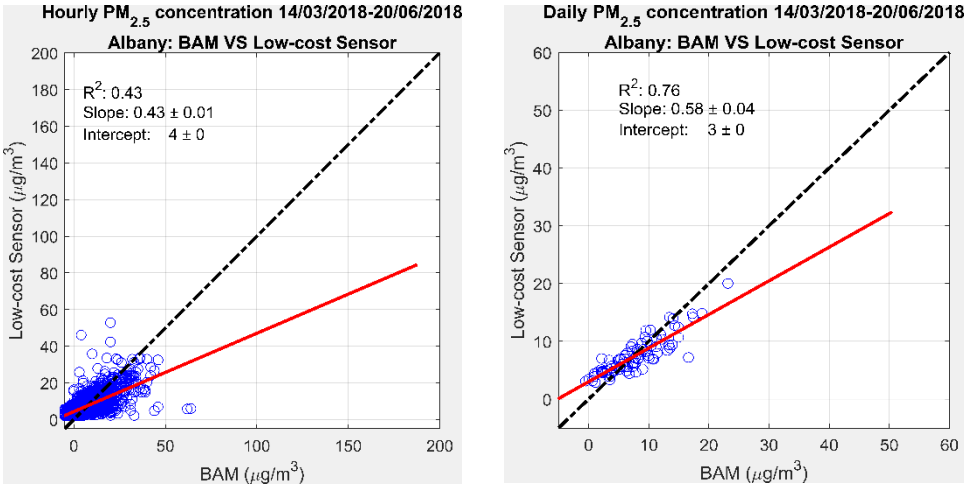

Figure C - 6 Comparison of hourly and daily $PM_{2.5}$ concentrations between the low-cost sensor (Albany) and BAM at the GA EPD site (Albany) from March 14, 2018 to June 20, 2018

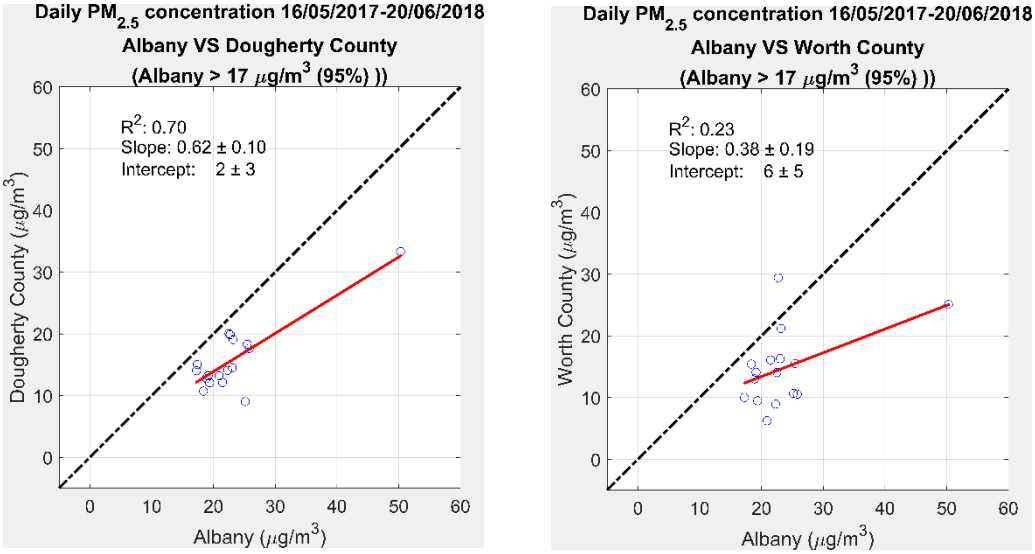

Figure C - 7 Comparison of daily $PM_{2.5}$ concentrations between high school low-cost sensors and BAM at GA EPD site (Albany) with a cut-off at 95% of BAM observation

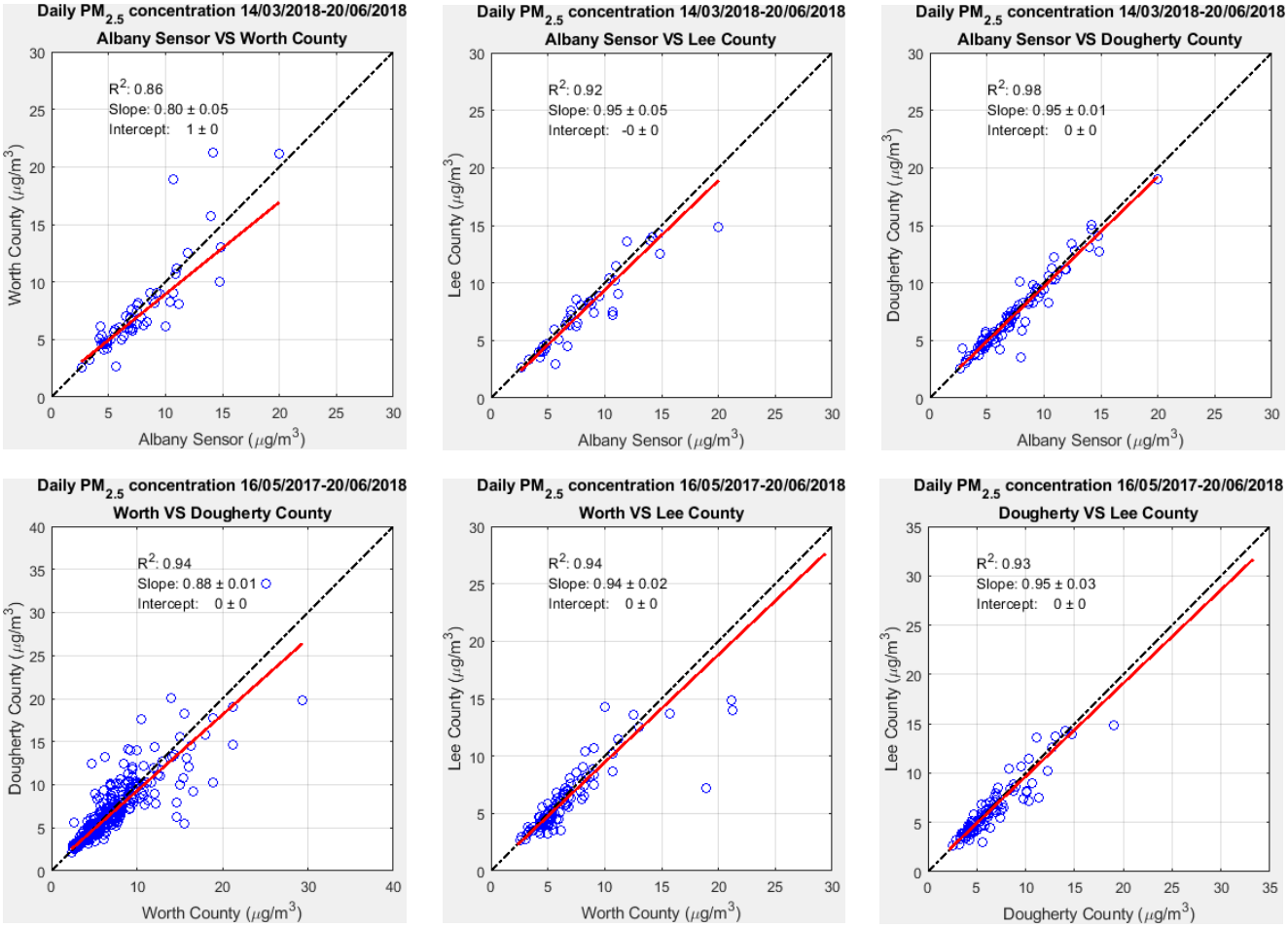

Figure C - 8 Orthogonal regression between low-cost sensors

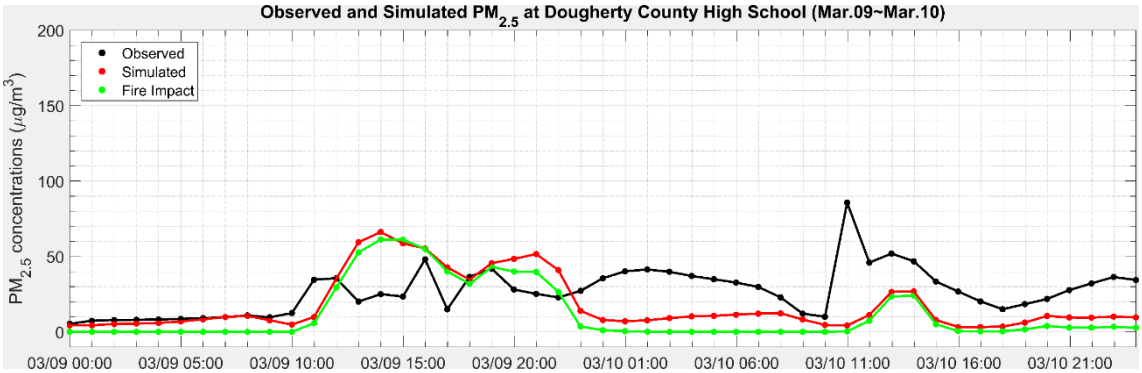

Figure C - 9 Comparison between observed and simulated $PM_{2.5}$ concentrations, and the fire impact from March 9 to March 10, 2018 at DCHS

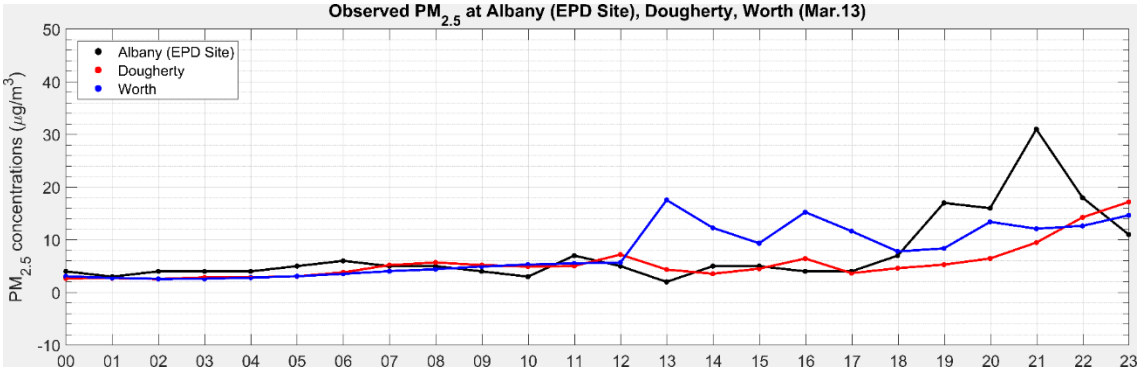

Figure C - 10 Observations on March 13 at GA EPD site (Albany), DCHS and WCHS

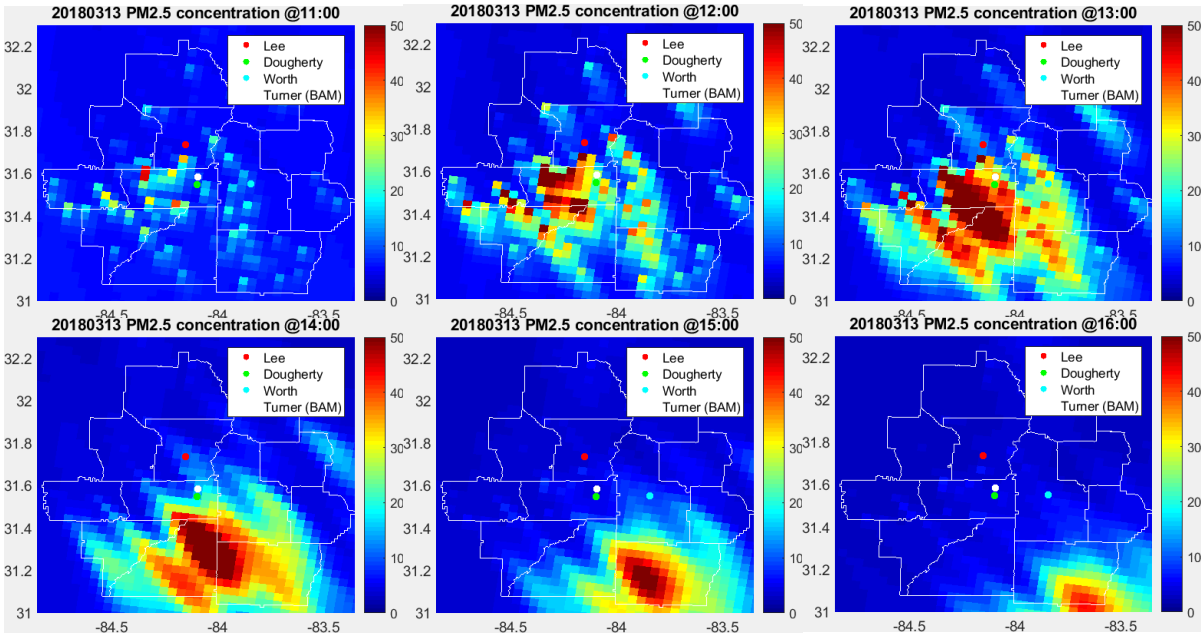

Figure C - 11 Spatial fields of hourly $PM_{2.5}$ concentration on March 13 from 11 a.m. to 4 p.m.: GA EPD site (Albany) (white dot), DCHS (green dot), WCHS (blue dot) and LCHS (red dot)

APPENDIX D. CHAPTER 5 SUPPLEMENTAL INFORMATION

A.1 **Data Fusion Method** (Friberg et al. 2016)

Annual mean fields $\bar{C}$ were calculated by using power regression analysis on the yearly mean of the daily observed concentrations ($\overline{OBS_m}$) at each monitor versus the mean of the CMAQ concentrations ($\overline{CMAQ_m}$) on corresponding days at each monitored grid location (Eq.1). The exponent β is constant for all years, while α changes annually (Table S1). Annual average CMAQ-derived fields are then constructed as:

$$\overline{\boldsymbol{C_{year}(s)}} = \boldsymbol{\alpha_{year}} \times \overline{\boldsymbol{CMAQ(s)}}^{\boldsymbol{\beta}} \quad (1)$$

Where, the overbar indicates annual temporal averaging, s corresponds to the space, and t represents the specific day being analyzed.

Daily interpolated observation fields, C_1, are created multiplying the annual mean CMAQ-derived fields ($\overline{\boldsymbol{C_{year}(s)}}$) by the interpolated daily ratios of daily observations normalized by annual mean observations at each monitor (Eq. 2). The interpolated fields were calculated using ordinary kriging method.

$$\boldsymbol{C_1(s,t)} = [\frac{\boldsymbol{OBS_m(t)}}{\overline{\boldsymbol{OBS_m}}}]_{\boldsymbol{krig}} \times \overline{\boldsymbol{C_{year}(s)}} \quad \textbf{(2)}$$

Daily CMAQ fields were scaled the annual average CMAQ-derived concentration field divided by the annual average of the CMAQ field at that location (Eq. 3).

$$\boldsymbol{C_2(s,t)} = \overline{\boldsymbol{C_{year}(s)}} \times [\frac{\boldsymbol{CMAQ(s,t)}}{\overline{\boldsymbol{CMAQ(s)}}}] \quad \textbf{(3)}$$

Combining C_1 and C_2 predicted fields via weighting the average based on daily estimation of error with following equations.

$$\boldsymbol{R_1(s,t) \approx R_{coll} e^{-D(s,t)\gamma}} \textbf{ (4)}$$

$$\boldsymbol{R_2 = \overline{R[OBS_i, CMAQ_i]}} \textbf{ (5)}$$

$$\boldsymbol{F(s,t) = \frac{R_1(\mathrm{s,t}) \times (1 - R_2)}{R_1(\mathrm{s,t}) \times (1 - R_2) + R_2 \times (1 - R_1(\mathrm{s,t}))}} \quad \textbf{(6)}$$

$$\boldsymbol{C^*(s,t) = F(s,t) \times C_1(\mathrm{s,t}) + (1 - F(s,t)) \times C_2(s,t)} \quad \textbf{(7)}$$

Where, R_1 is the estimated temporal correlation of observation as a function of distance from monitor, R_{coll} is the intercept which results from instrument error (i.e., error as estimated by collocated instruments), D is the distance between monitor locations, $\boldsymbol{\gamma}$ is the range at which the correlation between monitors has decreased to an *e-folding* of R_{coll} (Table D - 1), R_2 is the average of the temporal correlations at all monitors, *i* is the designation for each monitor and *F* is the weighting factor. C^* is the final fused field.

A.2 Integrated Mobile Source Indicator (IMSI) (Pachon et al. 2012)

IMSI uses the emission ratios of DV or GV for each species to separate the contribution ratios, and assume the temporal variations are the mobile source impact variations

$$IMSI = \frac{\frac{E_{EC,mobile}}{E_{EC,total}} \times C'_{EC} + \frac{E_{NO_X,mobile}}{E_{NO_X,total}} \times C'_{NO_X} + \frac{E_{CO,mobile}}{E_{CO,total}} \times C'_{CO}}{\frac{E_{EC,mobile}}{E_{EC,total}} + \frac{E_{NO_X,mobile}}{E_{NO_X,total}} + \frac{E_{CO,mobile}}{E_{CO,total}}} \quad (8)$$

$$IMSI_{DV} = \frac{\frac{E_{EC,DV}}{E_{EC,total}} \times C'_{EC} + \frac{E_{NO_X,DV}}{E_{NO_X,total}} \times C'_{NO_X} + \frac{E_{CO,DV}}{E_{CO,total}} \times C'_{CO}}{\frac{E_{EC,DV}}{E_{EC,total}} + \frac{E_{NO_X,DV}}{E_{NO_X,total}} + \frac{E_{CO,DV}}{E_{CO,total}}} \quad (9)$$

$$IMSI_{GV} = \frac{\frac{E_{EC,GV}}{E_{EC,total}} \times C'_{EC} + \frac{E_{NO_X,GV}}{E_{NO_X,total}} \times C'_{NO_X} + \frac{E_{CO,GV}}{E_{CO,total}} \times C'_{CO}}{\frac{E_{EC,GV}}{E_{EC,total}} + \frac{E_{NO_X,GV}}{E_{NO_X,total}} + \frac{E_{CO,GV}}{E_{CO,total}}} \quad (10)$$

where $IMSI$ (μg/m^3) represents the emission-based IMSI value, $IMSI_{GV}$ and $IMSI_{DV}$ represent the emission-based IMSI value for gasoline and diesel vehicles, $E_{EC,mobile}$ is EC emissions from mobile sources, $C_i' = C_i/\sigma_i$ denotes the normalized concentrations, and $C_i$and σ_i represent the average concentration and standard deviation of pollutant i in the period 2006 to 2008. Emissions for gasoline and diesel vehicles originated from the National Emissions Inventory (U.S. Environmental Protection Agency) and were processed using the Sparse Matrix Operator Kernel Emissions (SMOKE) model.

A.3 Two-stage Statistical Model (Hu et al. 2014a)

The stage-one model structure can be expressed as

$$PM_{2.5,st} = (b_0 + b_{0,t}) + (b_1 + b_{1,t})AOD_{st} + (b_2 + b_{2,t})\,\mathrm{Re}\,lativeHumidity_{st} + (b_3 + b_{3,t})WindSpeed_{st} + b_4 Elevation_s + b_5 MajorRoads_s + b_6 Forest\,Cover_s + b_7 Point\,Emissions_s + \varepsilon_{st}$$
$$(b_{0,t} b_{1,t} b_{2,t} b_{3,t}) \sim N[(0,0,0,0), \Psi]$$

where $PM_{2.5,st}$ is the measured ground level $PM_{2.5}$ concentration (μg/m^3) at site *s* on day *t*; b_0 and $b_{0,t}$ (day-specific) are the fixed and random intercept, respectively; AOD_{st} is the MODIS AOD value (unitless) at site *s* on day *t* ; b_1 and $b_{1,t}$ (day-specific) are the fixed and random slopes for AOD, respectively; *Relative Humidity*$_{st}$ is the relative humidity (%) at site *s* on day *t*; b_2 and $b_{2,t}$ (day-specific) are the fixed and random slopes for relative humidity, respectively; *Wind Speed*$_{st}$ is the 2-m wind speed (m/sec) at site *s* on day *t*; b_3

and $b_{3,t}$ (day-specific) are the fixed and random slopes for wind speed, respectively; $Elevation_s$ is elevation values (m) at site *s*; $Major\ Roads_s$ is road length values (m) at site *s*; $Forest\ Cover_s$ is forest cover values (unitless) at site *s*; $Point\ Emissions_s$ is point emissions (tons per year) at site *s*; and Ψ is an unstructured variance-covariance matrix for the random effects.

The stage-two model structure can be expressed as

$$PM_{2.5}\_resi_{st} = \beta_{0,s} + \beta_{1,s} AOD_{st} + \varepsilon_{st}$$

where $PM_{2.5}\_resi_{st}$ denotes the residuals from the stage one model at site *s* on day *t*, AOD_{st} is the MODIS AOD value (unitless) at site *s* on day *t*, and $\beta_{0,s}$ and $\beta_{1,s}$ are the location-specific intercept and slope.

A.4 **Neural Network-based Hybrid Model** (Di et al. 2016)

The hybrid prediction model was used to estimate $PM_{2.5}$ spatiotemporal fields. Briefly, data is used from satellite remote sensing (aerosol optical depth from the MODIS instrument on the Aqua and Terra satellites, absorbing aerosol index in the UV and visible range, column ozone, and NO_2 from the OMI instrument on the Aura satellite, and Normalized Difference Vegetation Index from MODIS instrument) for each available day from January 1st, 2000 to December 31st, 2012, with simulation outputs from a chemical transport model (GEOS-Chem, also run daily for that period), land-use terms (such as distance to major roads, emission sources, land use patterns, etc.), meteorological data (including temperature, humidity, wind speed, height of the planetary boundary layer, etc.) and other ancillary data to model monitored $PM_{2.5}$. A neural network approach

then combines these data to predict daily measurements at all EPA monitoring locations (N=1,928 for $PM_{2.5}$). Convolution layers capture spatial and temporal autocorrelations. To avoid overfitting, the optimal model was chosen based on prediction accuracy on left out monitors. The final model provided predictions of daily $PM_{2.5}$ levels in the continental United States with a spatial resolution of 1 km.

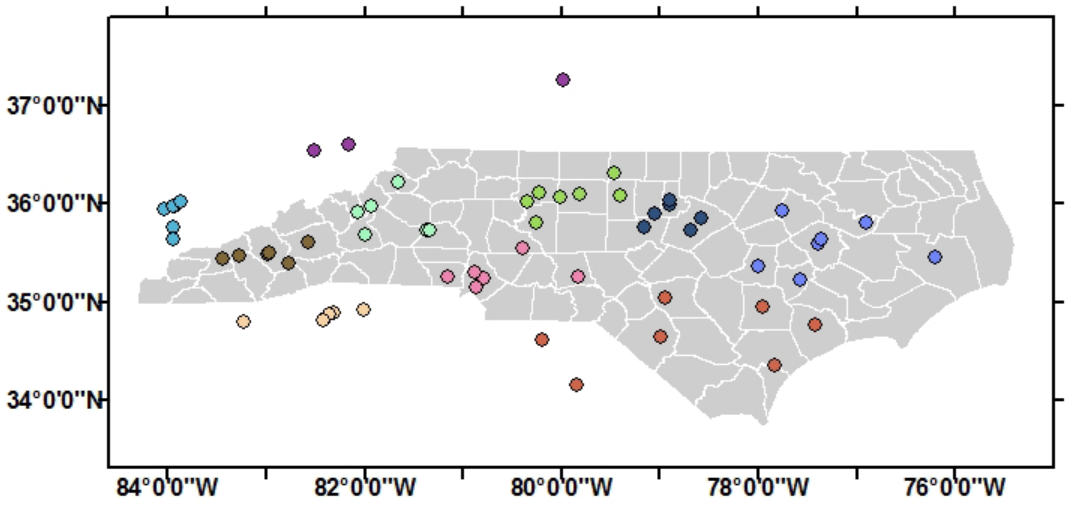

Figure D - 1 $PM_{2.5}$ monitor site (Each color represents a spatially-removed group)

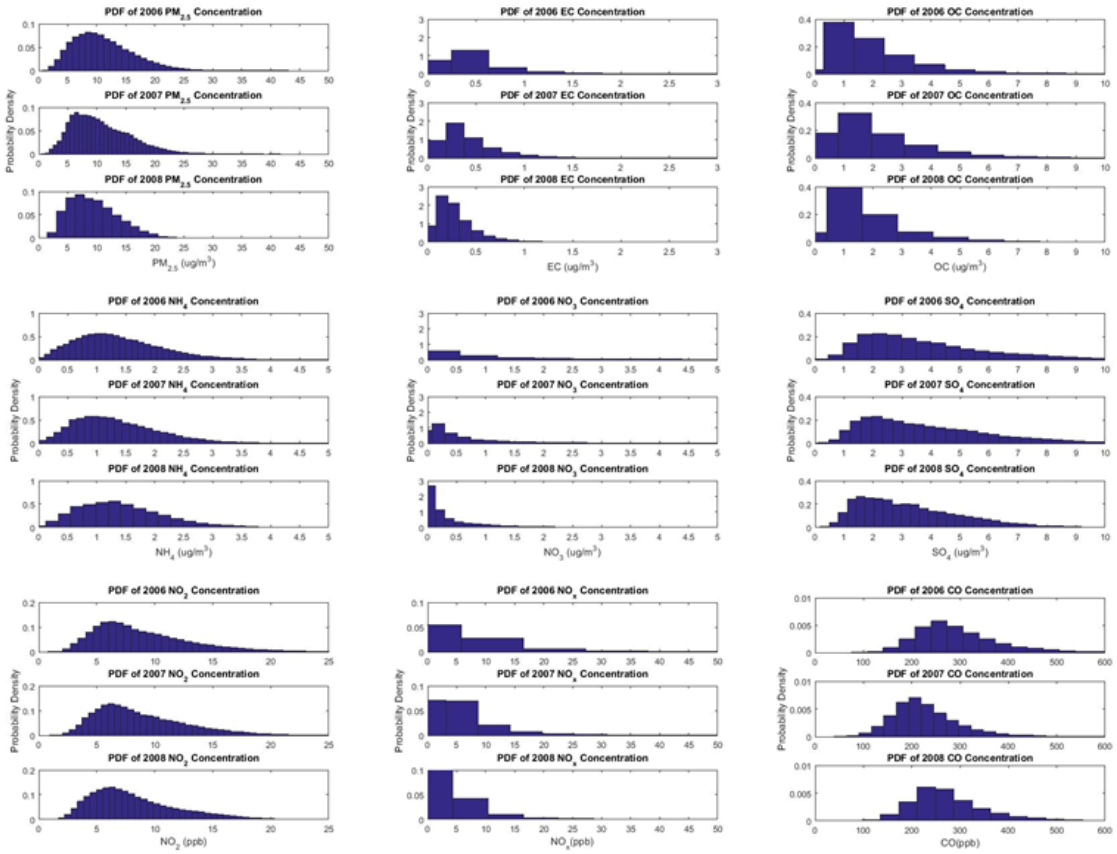

Figure D - 2 Probability density distribution of all species from 2006 to 2008

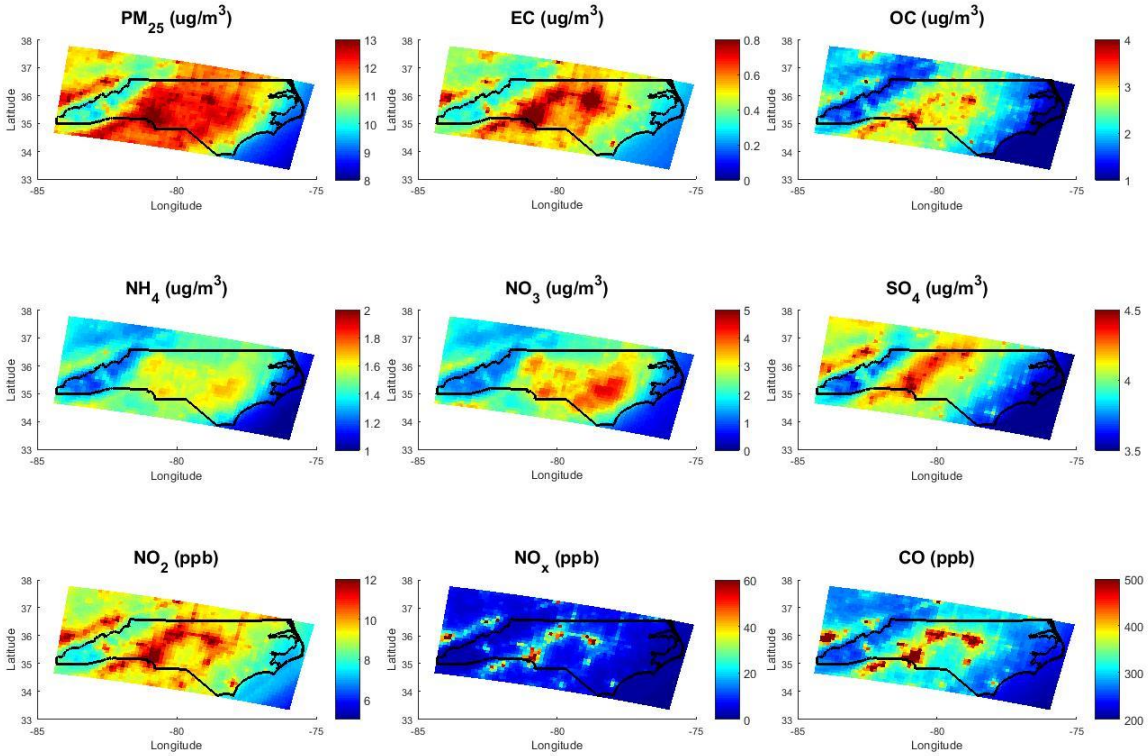

Figure D - 3 Annual average spatial distributions fields from Data Fusion, 2006

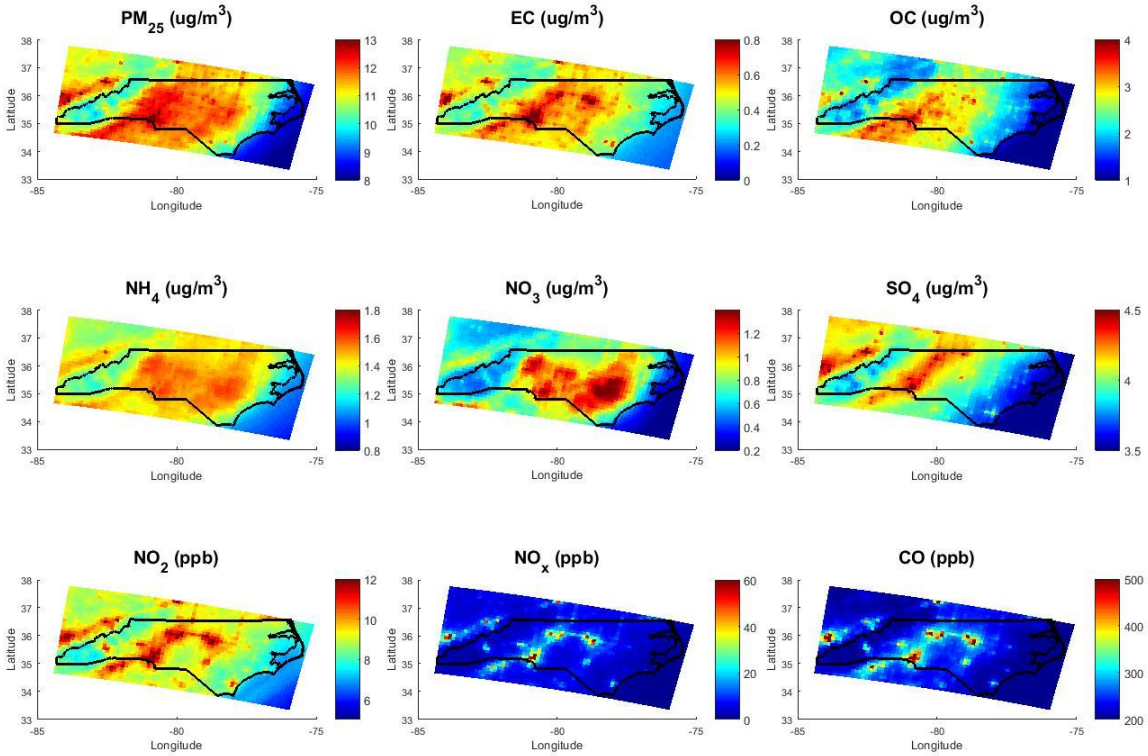

Figure D - 4 Annual average spatial distributions fields from Data Fusion, 2007

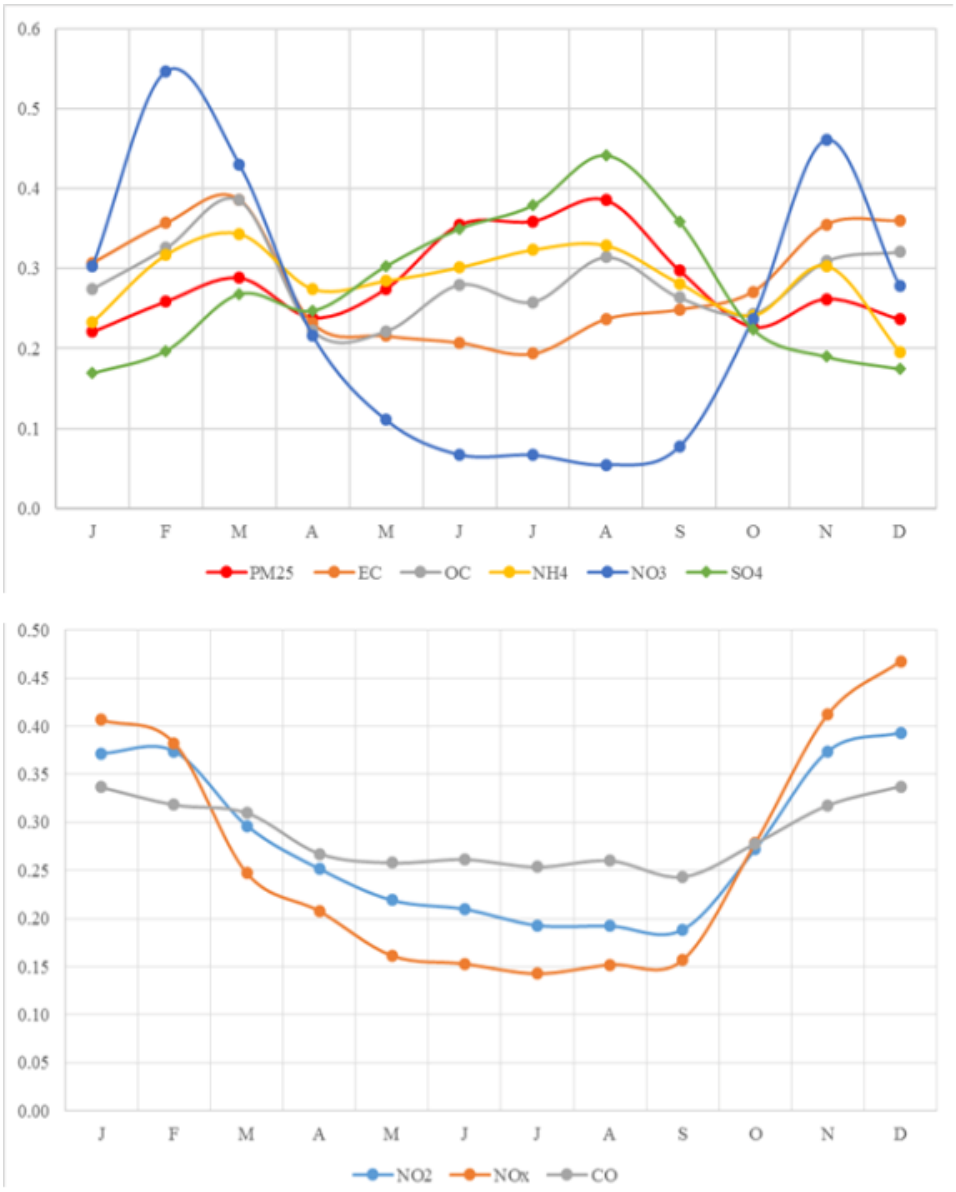

Figure D - 5 Normalized monthly average concentration for all species from 2006 to 2008

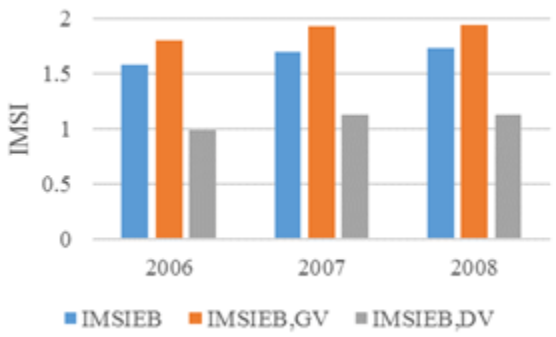

Figure D - 6 Annual trends of $IMSI_{EB}$, $IMSI_{EB, GV}$ and $IMSI_{EB, DV}$ from 2006 to 2008 (unitless)

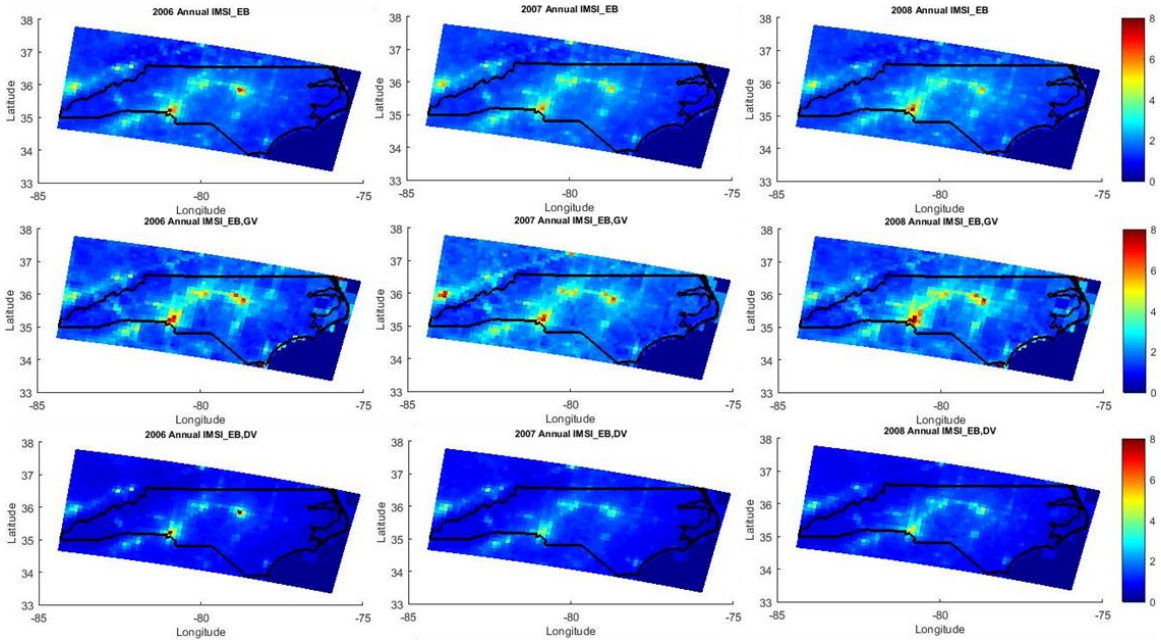

Figure D - 7 Annual $IMSI_{EB}$, $IMSI_{EB,\,GV}$ and $IMSI_{EB,\,DV}$ from 2006 to 2008

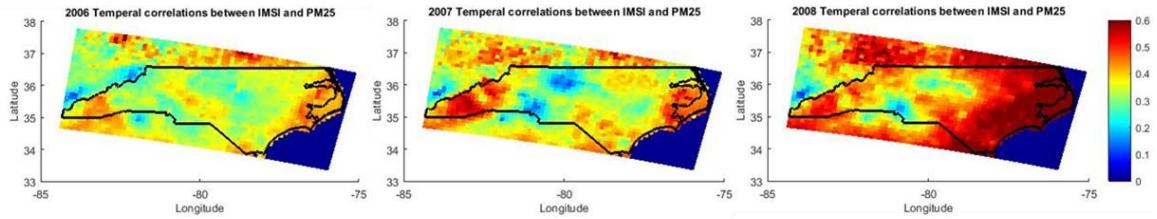

Figure D - 8 Temporal correlations between IMSI and $PM_{2.5}$ concentrations from 2006 to 2008

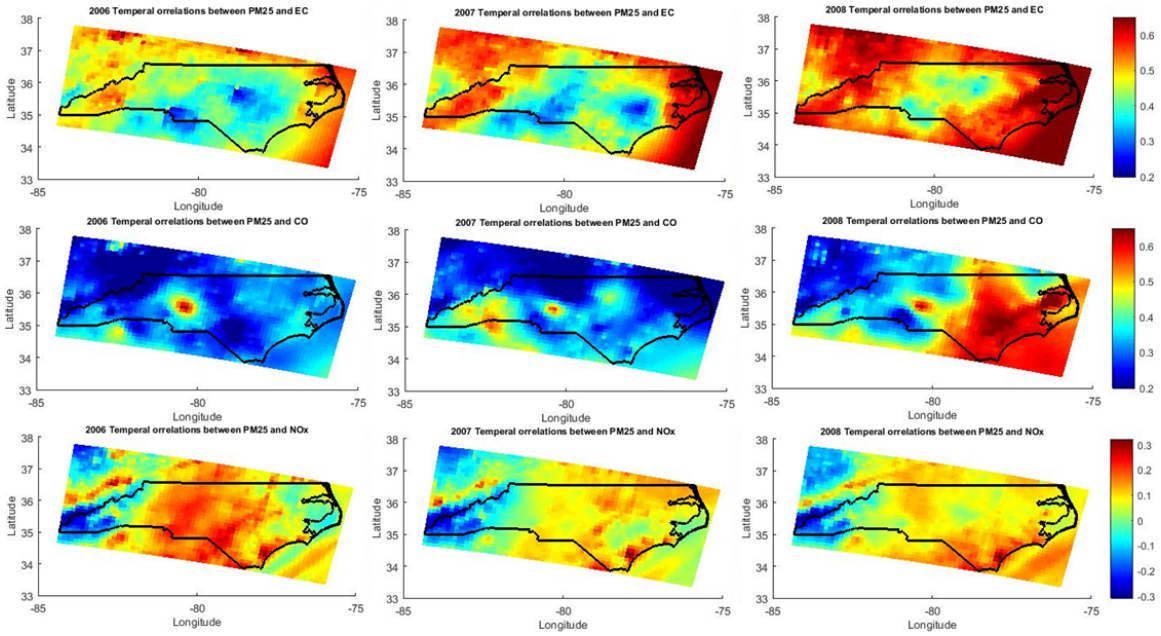

Figure D - 9 Temporal correlations between $PM_{2.5}$ and EC, CO, NO_x from 2006 to 2008

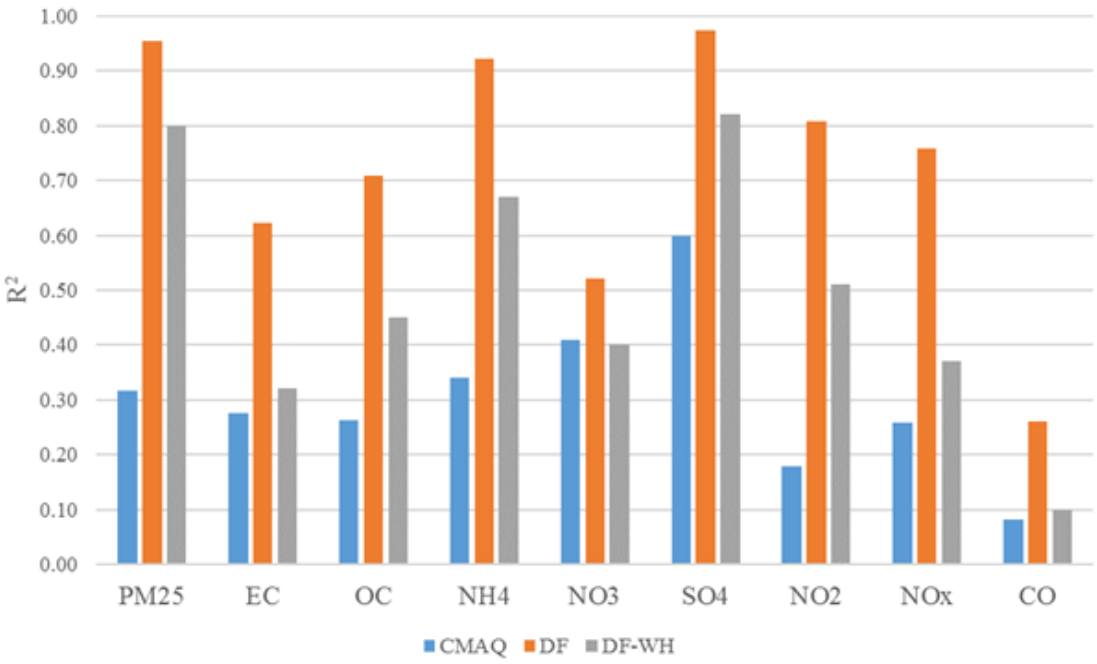

Figure D - 10 Comparison of R^2 between observations and simulated datasets (CMAQ, Data Fusion and 10% data-withheld Data Fusion) for 2006-2008

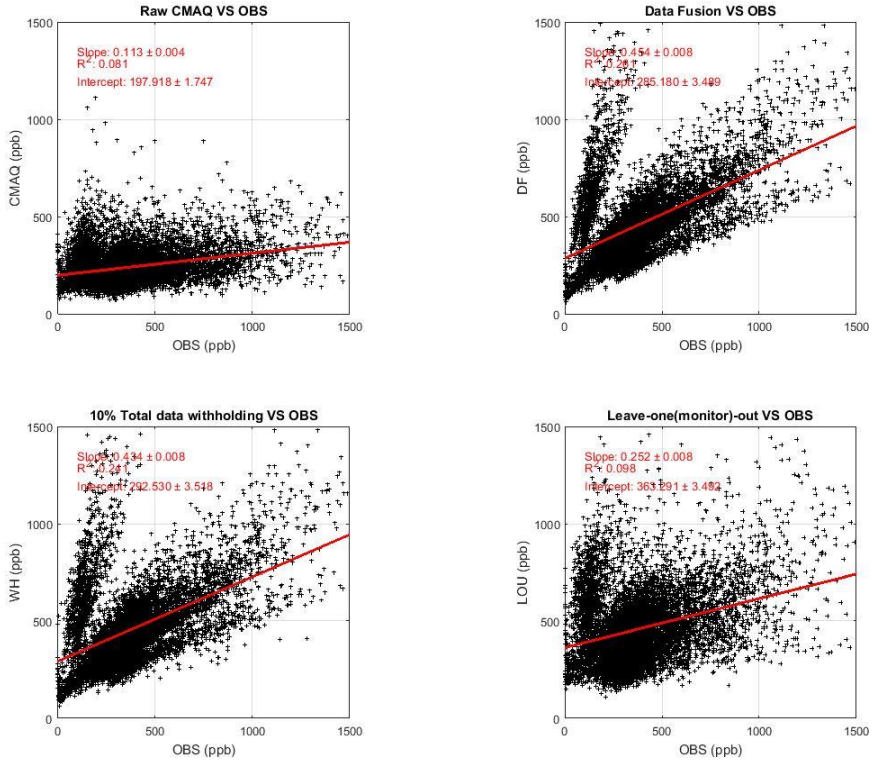

Figure D - 11 Linear regression between observation (OBS) and simulations (CO, data fusion)

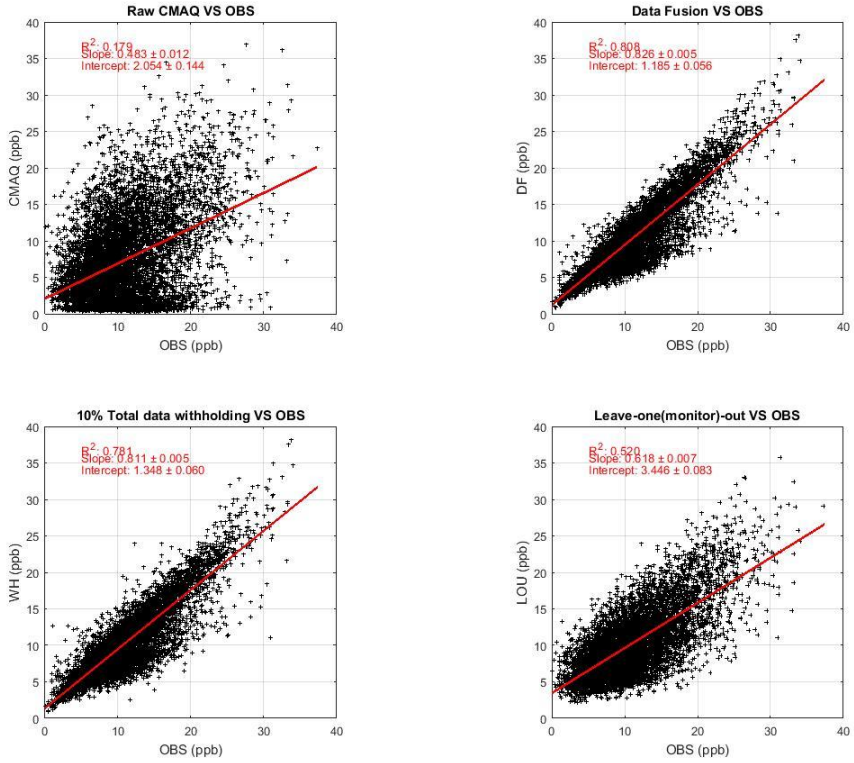

Figure D - 12Linear regression between observation (OBS) and simulations (NO_2)

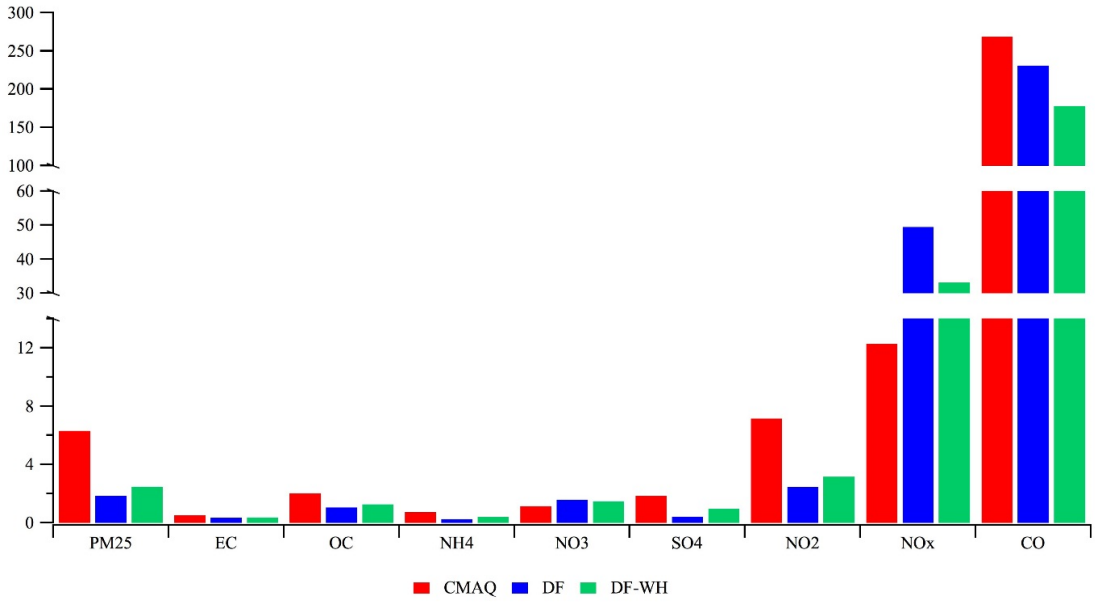

Figure D - 13 Comparison of RMSE between observations and simulated datasets (CMAQ, Data Fusion and 10% data-withheld Data Fusion) for 2006-2008 (ug/m^3:PM_{25}, EC, OC, NH_4^+, NO_3^-, SO_4^{2-}; ppb: NO2, NO_x, CO)

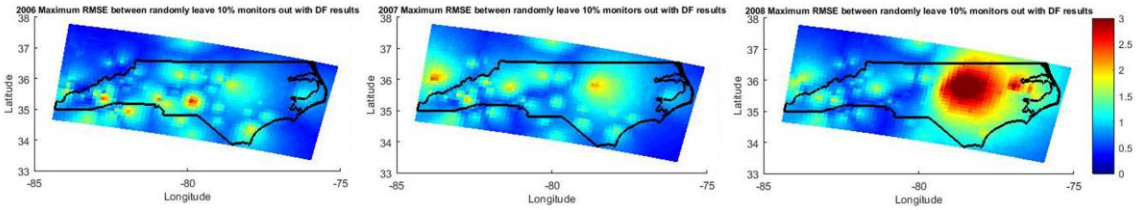

Figure D - 14 Maximum RMSD between leave-out-randomly (first time) and data fusion for all randomly leave 10% monitors out from 2006 (left) to 2008 (right)

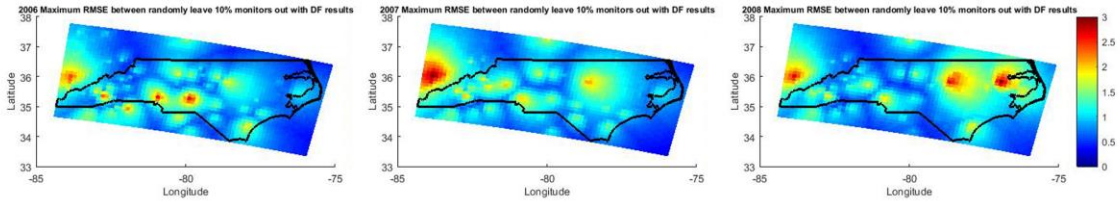

Figure D - 15 Maximum RMSD between leave-out-randomly (second time) and data fusion among all randomly leave 10% monitors out groups from 2006 (left) to 2008 (right).

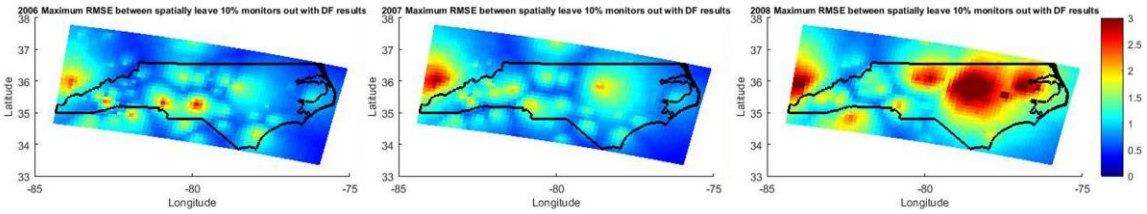

Figure D - 16 Maximum RMSD between leave-out-spatially and data fusion among all spatially leave out groups from 2006 (left) to 2008 (right)

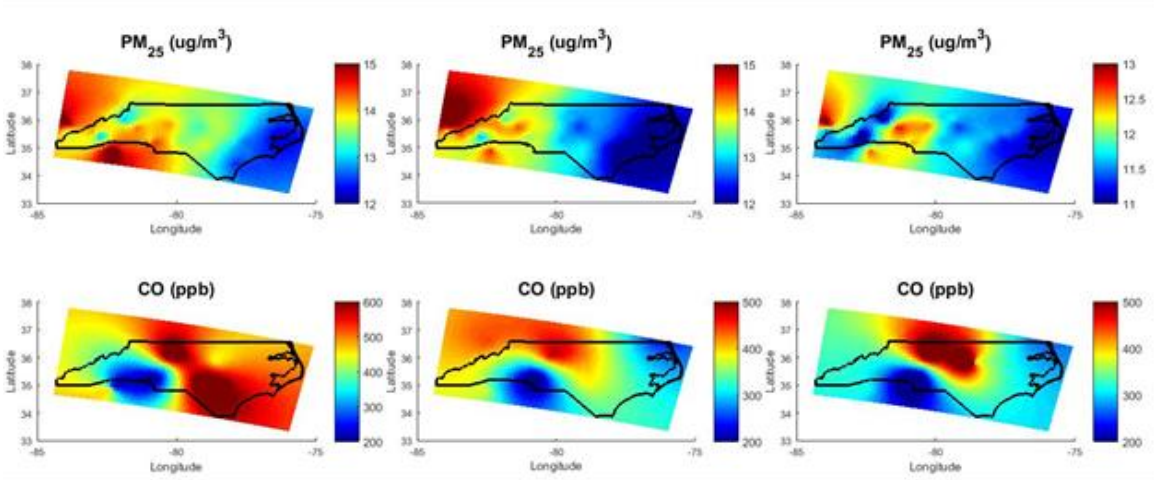

Figure D - 17 Annual average spatial distributions fields from Ordinary Kriging (2006, 2007, 2008)

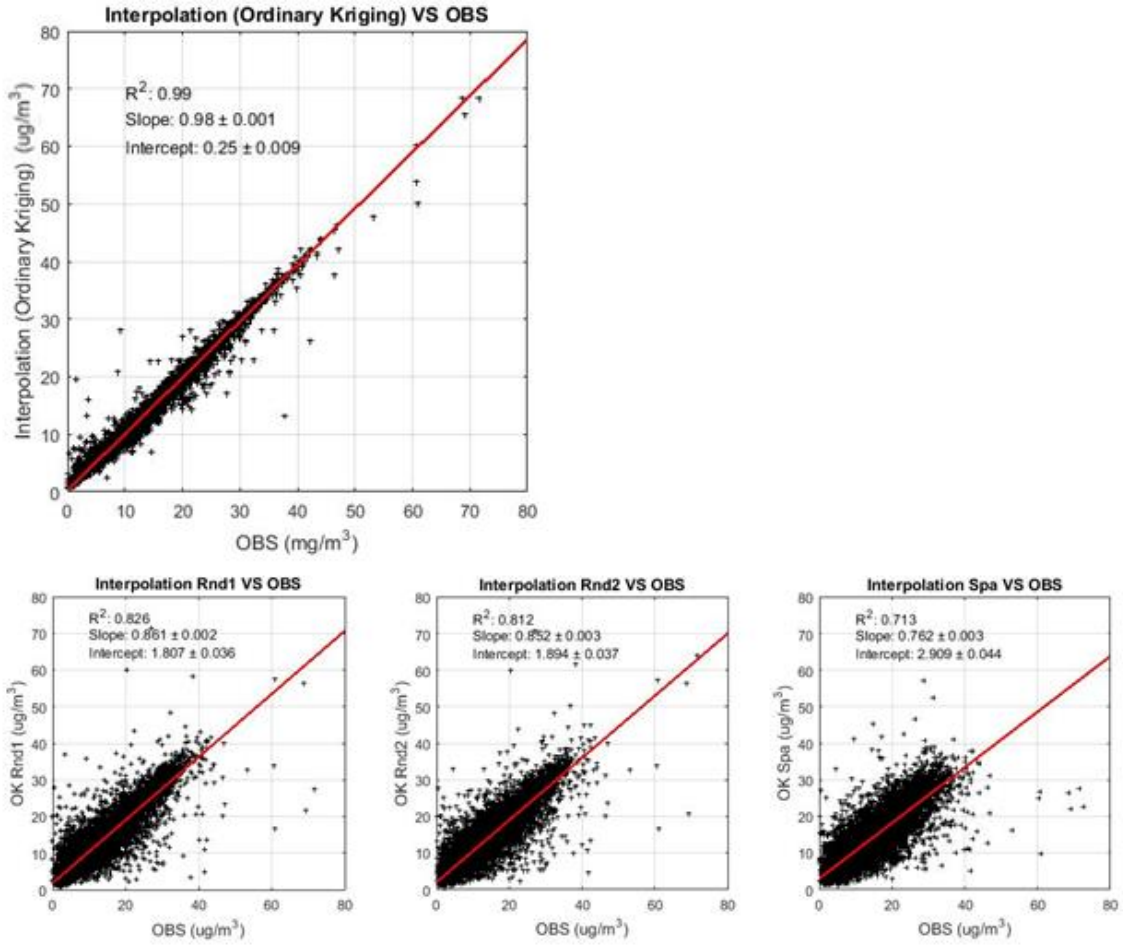

Figure D - 18 Linear regression between OBS and Ordinary kriging ($PM_{2.5}$, up: total data; done: Leave-monitors-out results)

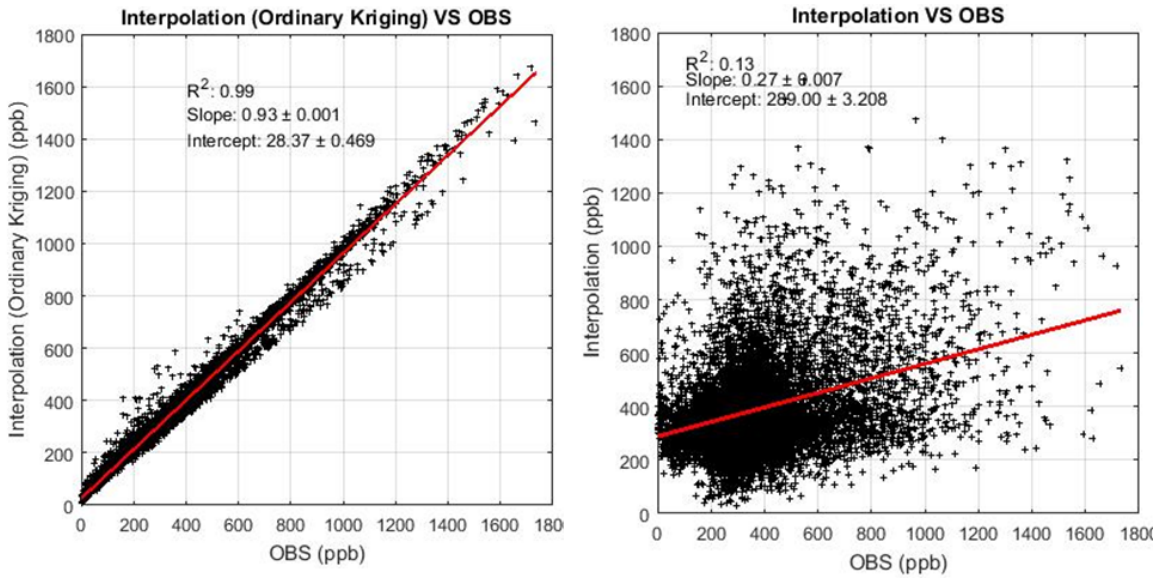

Figure D - 19 Linear regression between OBS and Ordinary kriging (CO, left: total data; right: Leave-one-out results)

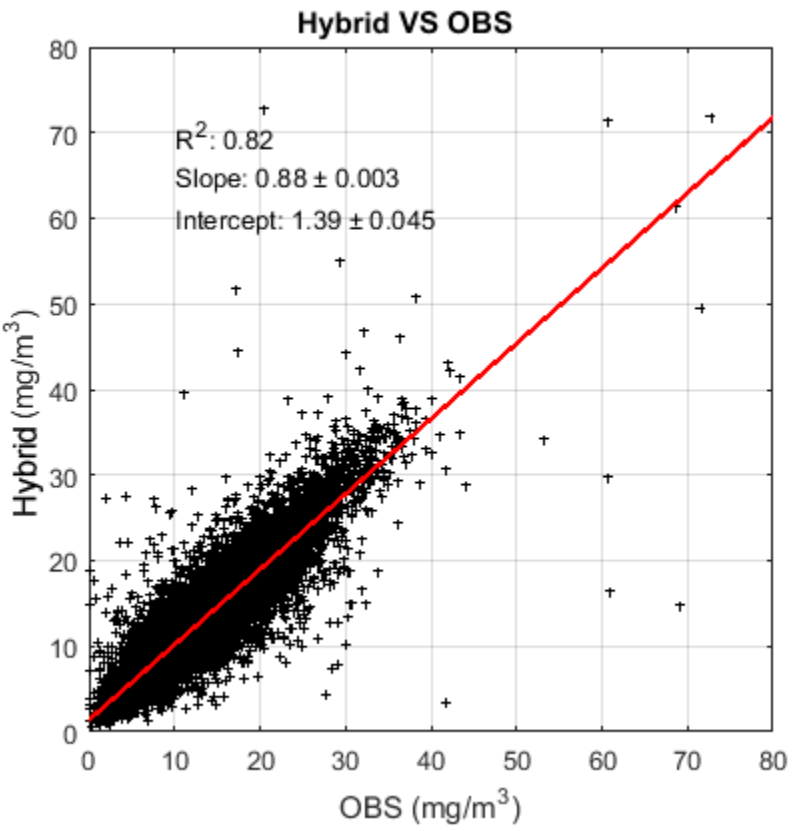

Figure D - 20 Linear regression between observation (OBS) and neural network-based hybrid model (Hybrid)

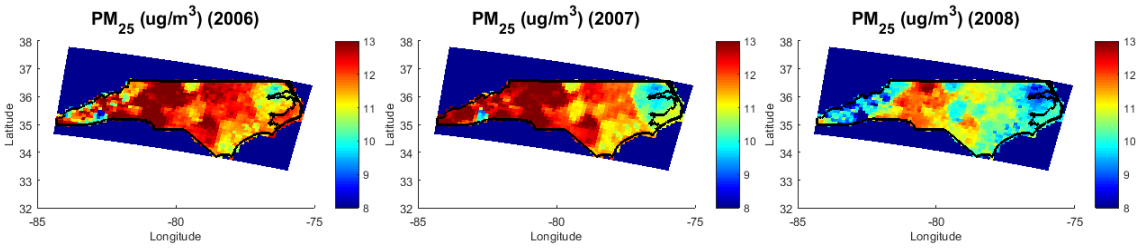

Figure D - 21 Annual average spatial distributions fields from neural network-based hybrid model for $PM_{2.5}$, 2006 – 2008 (12km)

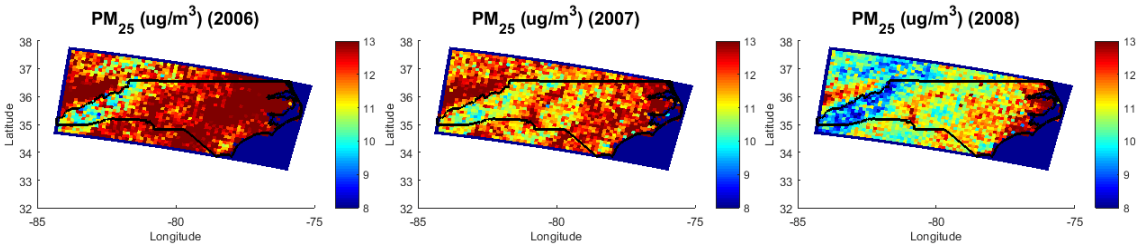

Figure D - 22 Annual average spatial distributions fields from two-stage statistical Model for $PM_{2.5}$, 2006 – 2008 (12km)

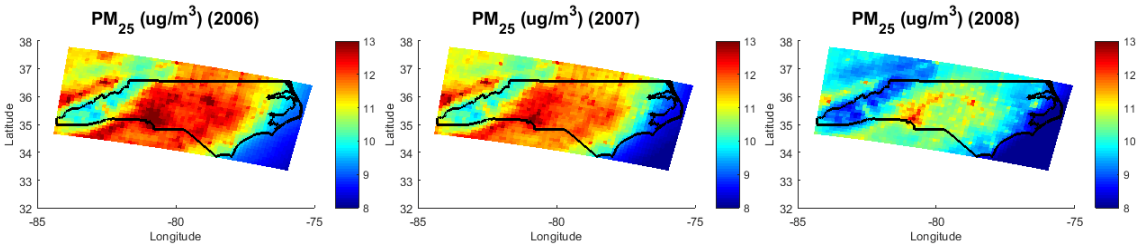

Figure D - 23 Annual average spatial distributions fields from data fusion for $PM_{2.5}$, 2006 – 2008 (12km)

Table D - 1 Parameters in Eq. 1 from 2006 to 2008 for 12-km resolution

		$PM_{2.5}$	EC	OC	NH_4	NO_3	SO_4	NO_2	NO_x	CO
< C* >	β	0.5	1	0.6	0.7	0.5	0.7	0.2	1.4	1.3
	α_{06}	4.2	1.4	1.7	1	0.5	1.7	7.50	1.7	0.4
	α_{07}	4.2	1.4	1.6	1	0.5	1.8	7.20	1.2	0.3
	α_{08}	3.6	1	1.5	1	0.4	1.8	6.90	1.0	0.4
R_2	average	0.53	0.45	0.46	0.51	0.57	0.71	0.44	0.49	0.40
R_1	R_{coll}	0.93	0.61	0.98	0.98	0.90	0.98	0.8	2.01	0.6
	γ	0.0022	0.0017	0.0030	0.0021	0.0018	0.0017	0.001	0.007	0.004

APPENDIX E. CHAPTER 6 SUPPLEMENTAL INFORMATION

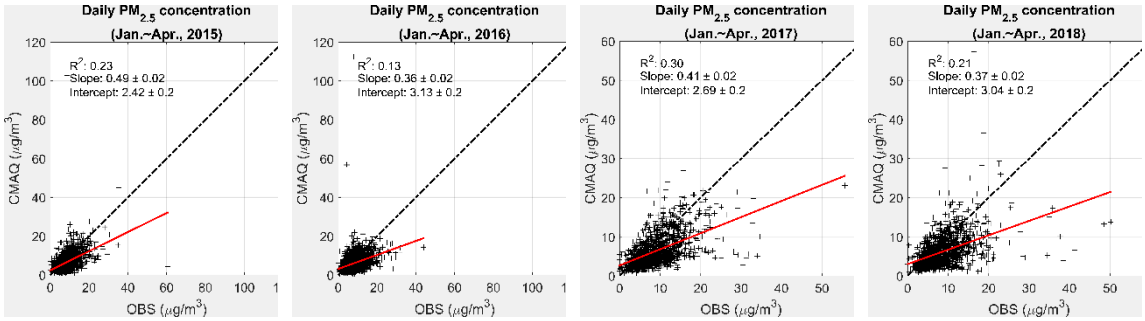

Figure E - 1 Comparison of daily total $PM_{2.5}$ concentration between observations (OBS) and CMAQ from 2015 to 2018, first four months

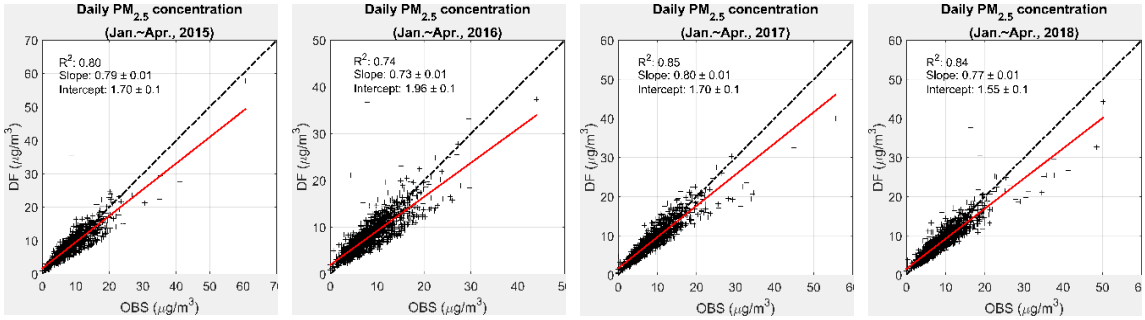

Figure E - 2 Comparison of daily total $PM_{2.5}$ concentration between observations (OBS) and DF from 2015 to 2018, first four months

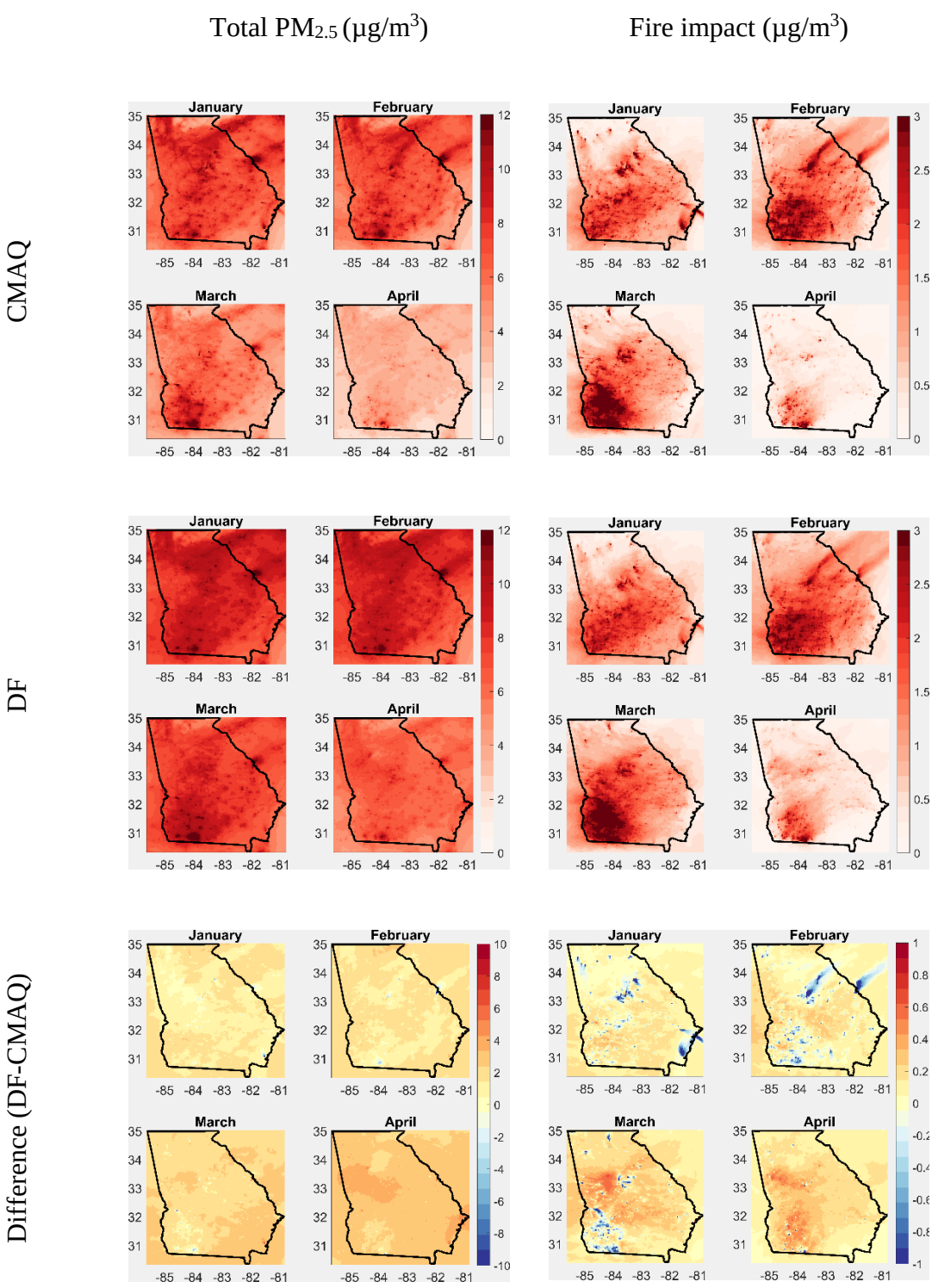

Figure E - 3 January – April monthly averages of total $PM_{2.5}$ and fire impact (2015): CMAQ-simulated, data fused and their difference

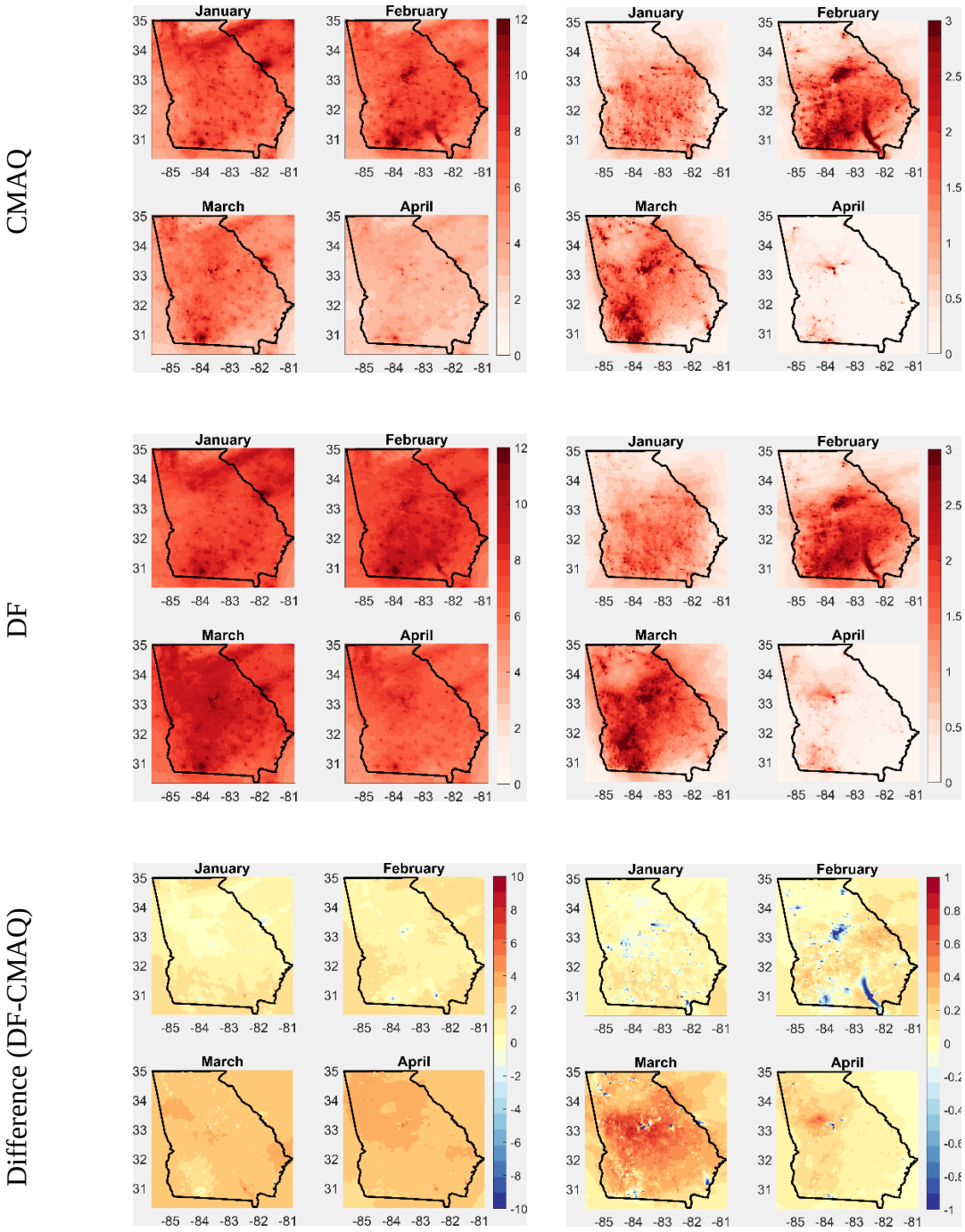

Figure E - 4 January – April monthly averages of total $PM_{2.5}$ and fire impact (2016): CMAQ-simulated, data fused and their difference

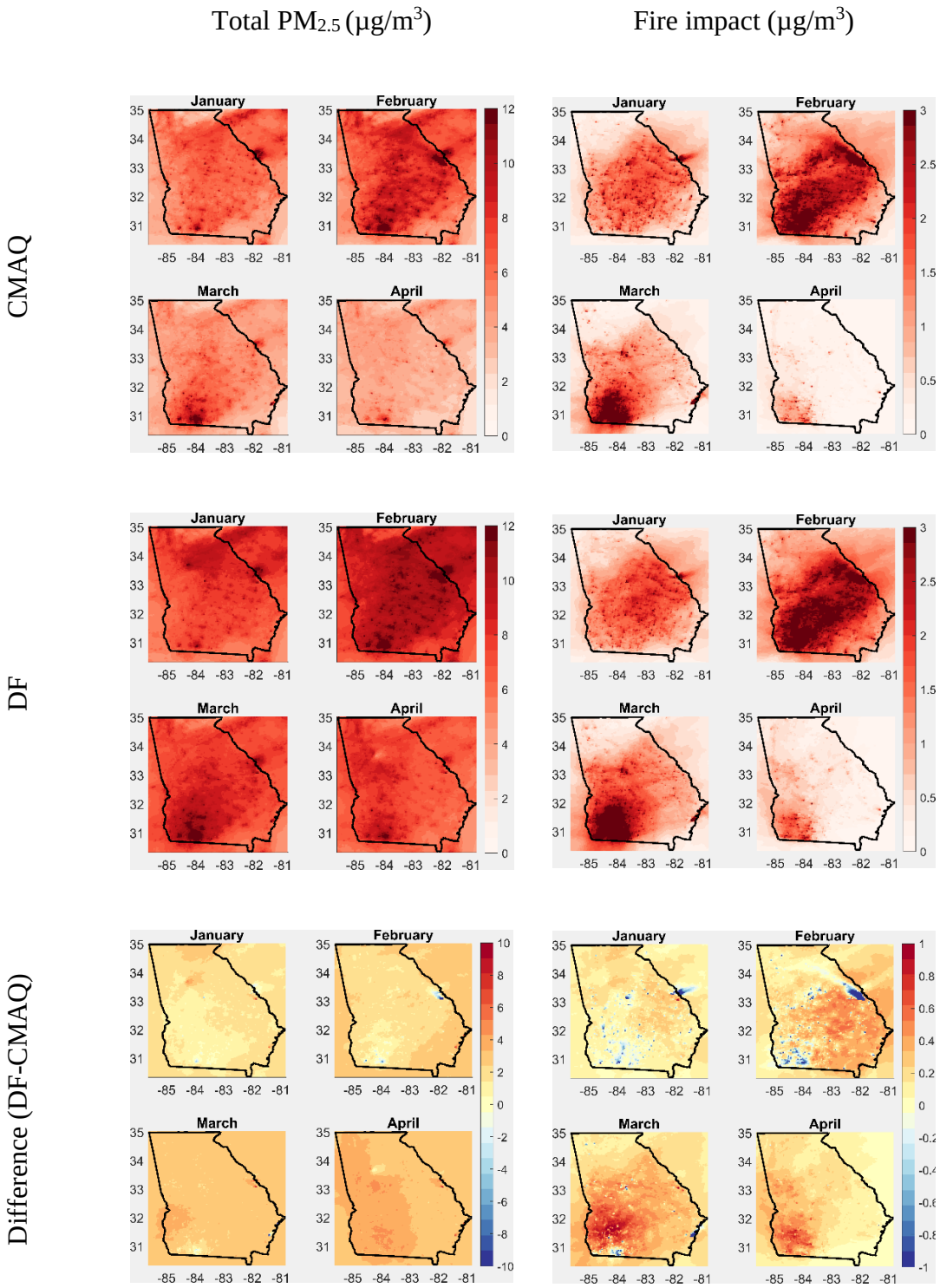

Figure E - 5 January – April monthly averages of total $PM_{2.5}$ and fire impact (2017): CMAQ-simulated, data fused and their difference

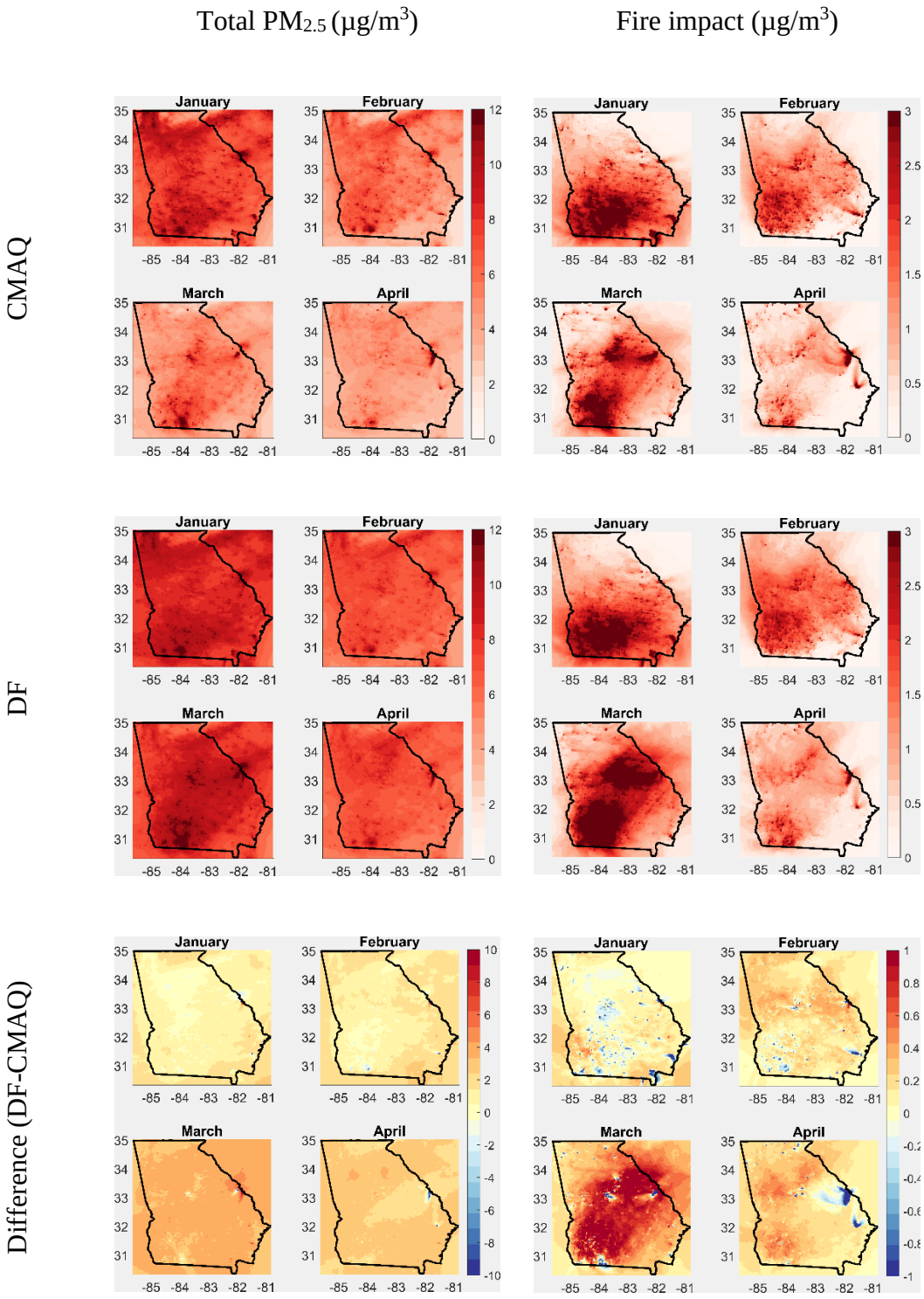

Figure E - 6 January – April monthly averages of total $PM_{2.5}$ and fire impact (2018): CMAQ-simulated, data fused and their difference

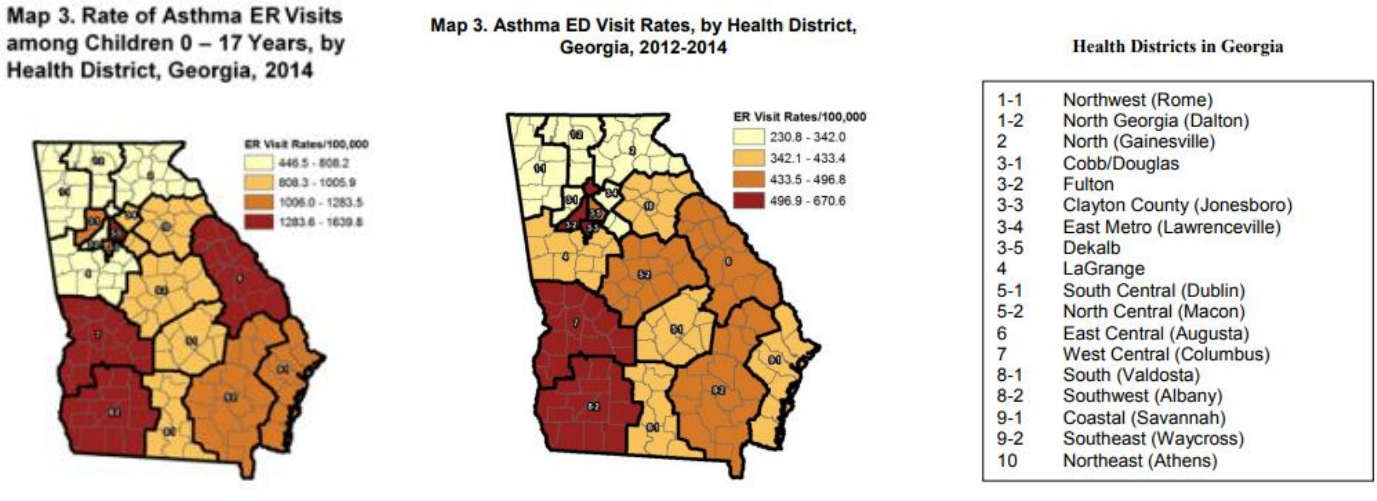

Figure E - 7 2016 Georgia Data Summary (asthma in children and adults) (Georgia Department of Public Health 2017a; Georgia Department of Public Health 2017b)

REFERENCES

Achtemeier GL, Goodrick SA, Liu Y (2012) Modeling Multiple-Core Updraft Plume Rise for an Aerial Ignition Prescribed Burn by Coupling Daysmoke with a Cellular Automata Fire Model. Atmosphere (Basel) 3:352–376. doi: 10.3390/atmos3030352

Alman BL, Pfister G, Hao H, et al (2016) The association of wildfire smoke with respiratory and cardiovascular emergency department visits in Colorado in 2012: a case crossover study. Environ Heal 15:64. doi: 10.1186/s12940-016-0146-8

Andreae MO, Merlet P (2001) Emission of trace gases and aerosols from biomass burning. Global Biogeochem Cycles 15:955–966. doi: 10.1029/2000GB001382

Baek J, Hu Y, Odman MT, Russell AG (2011) Modeling secondary organic aerosol in CMAQ using multigenerational oxidation of semi-volatile organic compounds. J Geophys Res Atmos 116:D22204. doi: 10.1029/2011JD015911

Balachandran S, Baumann K, Pachon JE, et al (2017) Evaluation of fire weather forecasts using PM2.5 sensitivity analysis. Atmos Environ 148:128–138. doi: 10.1016/J.ATMOSENV.2016.09.010

Barbero R, Abatzoglou JT, Larkin NK, et al (2015) Climate change presents increased potential for very large fires in the contiguous United States. Int J Wildl Fire. doi: 10.1071/WF15083

Beelen R, Hoek G, Pebesma E, et al (2009) Mapping of background air pollution at a fine spatial scale across the European Union. Sci Total Environ 407:1852–67. doi: 10.1016/j.scitotenv.2008.11.048

Bergin MS, Russell AG, Odman MT, et al (2008) Single-Source Impact Analysis Using Three-Dimensional Air Quality Models. J Air Waste Manage Assoc 58:1351–1359. doi: 10.3155/1047-3289.58.10.1351

Binkowski FS (2003) Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component 1. Model description. J Geophys Res 108:4183. doi: 10.1029/2001JD001409

Boschetti L, Eva HD, Brivio PA, Grégoire JM (2004) Lessons to be learned from the comparison of three satellite-derived biomass burning products. Geophys Res Lett 31:n/a–n/a. doi: 10.1029/2004GL021229

Brenner J, Goodrick S (2005) FLORIDA'S FIRE MANAGEMENT INFORMATION SYSTEM.

Brook RD, Franklin B, Cascio W, et al (2004) Air Pollution and Cardiovascular Disease.

Byun D, Schere KL (2006a) Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. Appl Mech Rev 59:51. doi: 10.1115/1.2128636

Byun D, Schere KL (2006b) Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. Appl Mech Rev 59:51. doi: 10.1115/1.2128636

Carlton AG, Turpin BJ, Altieri KE, et al (2008) CMAQ Model Performance Enhanced When In-Cloud Secondary Organic Aerosol is Included: Comparisons of Organic Carbon Predictions with Measurements. Environ Sci Technol 42:8798–8802. doi: 10.1021/es801192n

Carmona-Moreno C, Belward A, Malingreau J-P, et al (2005) Characterizing interannual variations in global fire calendar using data from Earth observing satellites. Glob Chang Biol 11:1537–1555. doi: 10.1111/j.1365-2486.2005.01003.x

Choi Y-J, Fernando HJS (2007) Simulation of smoke plumes from agricultural burns: Application to the San Luis/Rio Colorado airshed along the U.S./Mexico border. Sci Total Environ 388:270–289. doi: 10.1016/J.SCITOTENV.2007.07.058

Chu S-H (2004) PM2.5 episodes as observed in the speciation trends network. Atmos Environ 38:5237–5246. doi: 10.1016/j.atmosenv.2004.01.055

Chuvieco E, Yue C, Heil A, et al (2016) A new global burned area product for climate assessment of fire impacts. Glob Ecol Biogeogr 25:619–629. doi: 10.1111/geb.12440

Cohen AJ, Brauer M, Burnett R, et al (2017) Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet 389:1907–1918. doi: 10.1016/S0140-6736(17)30505-6

Connell BH, Gould KJ, Purdom JFW (2001) High-Resolution GOES-8 Visible and Infrared Cloud Frequency Composites over Northern Florida during the Summers 1996–99. Weather Forecast 16:713–724. doi: 10.1175/1520-0434(2001)016<0713:HRGVAI>2.0.CO;2

Cressie N (1988) Spatial prediction and ordinary kriging. Math Geol 20:405–421. doi: 10.1007/BF00892986

Darrow LA, Klein M, Flanders WD, et al (2014) Air Pollution and Acute Respiratory Infections Among Children 0–4 Years of Age: An 18-Year Time-Series Study. Am J Epidemiol 180:968–977. doi: 10.1093/aje/kwu234

Davis AY, Ottmar R, Liu Y, et al (2015) Fire emission uncertainties and their effect on smoke dispersion predictions: a case study at Eglin Air Force Base, Florida, USA. Int J Wildl Fire 24:276–285. doi: 10.1071/WF13071

Deming WE (1943) Statistical adjustment of data.

Dennekamp M, Abramson MJ (2011) The effects of bushfire smoke on respiratory health. Respirology 16:198–209. doi: 10.1111/j.1440-1843.2010.01868.x

Di Q, Kloog I, Koutrakis P, et al (2016) Assessing PM $_{2.5}$ Exposures with High Spatiotemporal Resolution across the Continental United States. Environ Sci Technol 50:4712–4721. doi: 10.1021/acs.est.5b06121

Dionisio KL, Baxter LK, Burke J, ?zkaynak H (2016) The importance of the exposure metric in air pollution epidemiology studies: When does it matter, and why? Air Qual Atmos Heal 9:495–502. doi: 10.1007/s11869-015-0356-1

Dohrenwend PB, Le M V, Bush JA, Thomas CF (2013) The impact on emergency department visits for respiratory illness during the southern california wildfires. West J Emerg Med 14:79–84. doi: 10.5811/westjem.2012.10.6917

Draxler RR, Hess GD (1997) NOAA Technical Memorandum ERL ARL-224 DESCRIPTION OF THE HYSPLIT_4 MODELING SYSTEM.

Draxler RR, Hess GD (1998) An overview of the HYSPLIT_4 modelling system for trajectories. Aust Meteorol Mag 47:295–308.

Dunker AM (1984) The decoupled direct method for calculating sensitivity coefficients in chemical kinetics. J Chem Phys 81:2385–2393. doi: 10.1063/1.447938

Dunker* AM, Yarwood G, Ortmann JP, Wilson GM (2002) Comparison of Source Apportionment and Source Sensitivity of Ozone in a Three-Dimensional Air Quality Model. doi: 10.1021/ES011418F

Emery C, Liu Z, Russell AG, et al (2017) Recommendations on statistics and benchmarks

to assess photochemical model performance. J Air Waste Manage Assoc 67:582–598. doi: 10.1080/10962247.2016.1265027

Environ R (2016) User's Guide—Comprehensive Air-Quality Model with Extensions, Version 6.3.

EPA U (2011) 2011 National Emissions Inventory (NEI) Documentation.

Fann N, Alman B, Broome RA, et al (2018) The health impacts and economic value of wildland fire episodes in the US: 2008-2012. Sci Total Environ 610:802–809. doi: 10.1016/j.scitotenv.2017.08.024

Faustini A, Alessandrini ER, Pey J, et al (2015) Short-term effects of particulate matter on mortality during forest fires in Southern Europe: results of the MED-PARTICLES Project. Occup Environ Med 72:323–9. doi: 10.1136/oemed-2014-102459

Forouzanfar MH, Afshin A, Alexander LT, et al (2016) Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 388:1659–1724. doi: 10.1016/S0140-6736(16)31679-8

Fraser RH, Li Z, Cihlar J (2000a) Hotspot and NDVI Differencing Synergy (HANDS) A New Technique for Burned Area Mapping over Boreal Forest. Remote Sens Environ 74:362–376. doi: 10.1016/S0034-4257(00)00078-X

Fraser RH, Li Z, Landry R (2000b) SPOT VEGETATION for characterizing boreal forest fires. Int J Remote Sens 21:3525–3532. doi: 10.1080/014311600750037534

Friberg MD, Zhai X, Holmes HA, et al (2016) Method for Fusing Observational Data and Chemical Transport Model Simulations To Estimate Spatiotemporally Resolved Ambient Air Pollution. Environ Sci Technol 50:3695–3705. doi: 10.1021/acs.est.5b05134

Gao M, Cao J, Seto E (2015) A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi'an, China. Environ Pollut 199:56–65. doi: 10.1016/J.ENVPOL.2015.01.013

Garcia-Menendez F, Hu Y, Odman MT (2014) Simulating smoke transport from wildland fires with a regional-scale air quality model: Sensitivity to spatiotemporal allocation of fire emissions. Sci Total Environ 493:544–553. doi: 10.1016/J.SCITOTENV.2014.05.108

Garcia-Menendez F, Hu Y, Odman MT (2013) Simulating smoke transport from wildland fires with a regional-scale air quality model: Sensitivity to uncertain wind fields. J Geophys Res Atmos 118:6493–6504. doi: 10.1002/jgrd.50524

Georgia Department of Public Health (2017a) 2016 GEORGIA DATA SUMMARY | ADULT ASTHMA.

Georgia Department of Public Health (2017b) 2016 GEORGIA DATA SUMMARY | ASTHMA IN CHILDREN.

Gertler AW (2005) Diesel vs. gasoline emissions: Does PM from diesel or gasoline vehicles dominate in the US? Atmos Environ 39:2349–2355. doi: 10.1016/j.atmosenv.2004.05.065

Gertler AW, Gillies JA, Pierson WR An Assessment of the Mobile Source Contribution to PM10 and PM2.5 in the United States. Water Air Soil Pollut 123:203–214. doi: 10.1023/A:1005263220659

Gibson HM, Vonder Haar TH (1990) Cloud and Convection Frequencies over the Southeast United States as Related to Small-Scale Geographic Features. Mon Weather Rev 118:2215–2227. doi: 10.1175/1520-0493(1990)118<2215:CACFOT>2.0.CO;2

Giglio L, Descloitres J, Justice CO, Kaufman YJ (2003) An Enhanced Contextual Fire Detection Algorithm for MODIS. Remote Sens Environ 87:273–282. doi: 10.1016/S0034-4257(03)00184-6

Giglio L, Randerson JT, van der Werf GR, et al (2010) Assessing variability and long-term trends in burned area by merging multiple satellite fire products. Biogeosciences 7:1171–1186. doi: 10.5194/bg-7-1171-2010

Giglio L, Randerson JT, van der Werf GR (2013) Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). J Geophys Res Biogeosciences 118:317–328. doi: 10.1002/jgrg.20042

Giglio L, Schroeder W, Justice CO (2016) The collection 6 MODIS active fire detection algorithm and fire products. Remote Sens Environ 178:31–41. doi: 10.1016/J.RSE.2016.02.054

Gilboa SM, Mendola P, Olshan AF, et al (2005) Relation between ambient air quality and selected birth defects, seven county study, Texas, 1997-2000. Am J Epidemiol 162:238–52. doi: 10.1093/aje/kwi189

Gilliland AB, Hogrefe C, Pinder RW, et al (2008) Dynamic evaluation of regional air quality models: Assessing changes in O3 stemming from changes in emissions and meteorology. Atmos Environ 42:5110–5123. doi: 10.1016/j.atmosenv.2008.02.018

Godowitch JM, Gilliam RC, Roselle SJ (2015) Investigating the impact on modeled ozone concentrations using meteorological fields from WRF with an updated four–dimensional data assimilation approach. Atmos Pollut Res 6:305–311. doi: 10.5094/APR.2015.034

Goodrick SL, Achtemeier GL, Larkin NK, et al (2013) Modelling smoke transport from wildland fires: a review. Int J Wildl Fire 22:83. doi: 10.1071/WF11116

Gupta P, Doraiswamy P, Levy R, et al (2018) Impact of California Fires on Local and Regional Air Quality: The Role of a Low-Cost Sensor Network and Satellite Observations. GeoHealth. doi: 10.1029/2018GH000136

Haikerwal A, Akram M, Del Monaco A, et al (2015) Impact of Fine Particulate Matter (PM2.5) Exposure During Wildfires on Cardiovascular Health Outcomes. J Am Heart Assoc 4:e001653. doi: 10.1161/JAHA.114.001653

Hakami A, M. Talat Odman;, Russell AG (2003) High-Order, Direct Sensitivity Analysis of Multidimensional Air Quality Models. doi: 10.1021/ES020677H

Han I, Symanski E, Stock TH (2017) Feasibility of using low-cost portable particle monitors for measurement of fine and coarse particulate matter in urban ambient air. J Air Waste Manage Assoc 67:330–340. doi: 10.1080/10962247.2016.1241195

Hawbaker TJ, Vanderhoof MK, Beal YG, et al (2017) Landsat Burned Area Essential Climate Variable products for the conterminous United States (1984 -2015). US Geol Surv data release. doi: 10.5066/F73B5X76

He L, Li Z (2012) Enhancement of a fire detection algorithm by eliminating solar reflection in the mid-IR band: application to AVHRR data. Int J Remote Sens 33:7047–7059. doi: 10.1080/2150704X.2012.699202

Hoelzemann JJ (2004) Global Wildland Fire Emission Model (GWEM): Evaluating the use of global area burnt satellite data. J Geophys Res 109:D14S04. doi: 10.1029/2003JD003666

Hu X, Waller LA, Al-Hamdan MZ, et al (2013) Estimating ground-level PM(2.5) concentrations in the southeastern U.S. using geographically weighted regression. Environ Res 121:1–10. doi: 10.1016/j.envres.2012.11.003

Hu X, Waller LA, Lyapustin A, et al (2014a) Estimating ground-level PM2.5 concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model. Remote Sens Environ 140:220–232. doi: 10.1016/j.rse.2013.08.032

Hu X, Yu C, Tian D, et al (2016) Comparison of the Hazard Mapping System (HMS) fire product to ground-based fire records in Georgia, USA. J Geophys Res Atmos 121:2901–2910. doi: 10.1002/2015JD024448

Hu Y, Balachandran S, Pachon JE, et al (2014b) Fine particulate matter source apportionment using a hybrid chemical transport and receptor model approach. Atmos Chem Phys 14:5415–5431. doi: 10.5194/acp-14-5415-2014

Hu Y, Odman MT, Chang ME, Russell AG (2015) Operational forecasting of source impacts for dynamic air quality management. Atmos Environ 116:320–322. doi: 10.1016/J.ATMOSENV.2015.04.061

Huang R, Zhai X, Ivey CE, et al (2018a) Air pollutant exposure field modeling using air quality model-data fusion methods and comparison with satellite AOD-derived fields: application over North Carolina, USA. Air Qual Atmos Heal 11:11–22. doi: 10.1007/s11869-017-0511-y

Huang R, Zhang X, Chan D, et al (2018b) Burned Area Comparisons Between Prescribed Burning Permits in Southeastern United States and Two Satellite-Derived Products. J Geophys Res Atmos 123:4746–4757. doi: 10.1029/2017JD028217

Hubbell B (2012) Understanding urban exposure environments: new research directions for informing implementation of U.S. air quality standards. Air Qual Atmos Heal 5:259–267. doi: 10.1007/s11869-011-0153-4

In Statista - The Statistics Portal. (2018) National Interagency Fire Center (National Interagency Coordination Center). (n.d.). Number of fires and acres burned due to U.S. prescribed fires in 2017, by state. In: Stat. - Stat. Portal. https://www.statista.com/statistics/204014/highest-number-of-prescribed-fires-in-the-us-by-states/. Accessed 6 Aug 2018

Ivey CE, Holmes HA, Hu Y, et al (2016) A method for quantifying bias in modeled concentrations and source impacts for secondary particulate matter. Front Environ Sci Eng 10:14. doi: 10.1007/s11783-016-0866-6

Ivey CE, Holmes HA, Hu YT, et al (2015) Development of PM2.5 source impact spatial fields using a hybrid source apportionment air quality model. Geosci Model Dev

8:2153–2165. doi: 10.5194/gmd-8-2153-2015

Jathar SH, Cappa CD, Wexler AS, et al (2016) Simulating secondary organic aerosol in a regional air quality model using the statistical oxidation model – Part 1: Assessing the influence of constrained multi-generational ageing. Atmos Chem Phys 16:2309–2322. doi: 10.5194/acp-16-2309-2016

Johnson KK, Bergin MH, Russell AG, Hagler GSW (2018) Field Test of Several Low-Cost Particulate Matter Sensors in High and Low Concentration Urban Environments. Aerosol Air Qual Res 18:565–578. doi: 10.4209/aaqr.2017.10.0418

Johnson KK, Bergin MH, Russell AG, Hagler GSW (2016) Using Low Cost Sensors to Measure Ambient Particulate Matter Concentrations and On-Road Emissions Factors. Atmos Meas Tech Discuss 1–22. doi: 10.5194/amt-2015-331

Johnson M, Isakov V, Touma JS, et al (2010) Evaluation of land-use regression models used to predict air quality concentrations in an urban area. Atmos Environ 44:3660–3668. doi: 10.1016/j.atmosenv.2010.06.041

Johnston FH, Purdie S, Jalaludin B, et al (2014) Air pollution events from forest fires and emergency department attendances in Sydney, Australia 1996–2007: a case-crossover analysis. Environ Heal 13:105. doi: 10.1186/1476-069X-13-105

Joint Fire Science Program (2009) Consume 3.0-A Software Tool for Computing Fuel Consumption.

Jovašević-Stojanović M, Bartonova A, Topalović D, et al (2015) On the use of small and cheaper sensors and devices for indicative citizen-based monitoring of respirable particulate matter. Environ Pollut 206:696–704. doi: 10.1016/J.ENVPOL.2015.08.035

Kanaroglou PS, Jerrett M, Morrison J, et al (2005) Establishing an air pollution monitoring network for intra-urban population exposure assessment: A location-allocation approach. Atmos Environ 39:2399–2409. doi: 10.1016/j.atmosenv.2004.06.049

Kelleher S, Quinn C, Miller-Lionberg D, Volckens J (2018) A low-cost particulate matter (PM2.5) monitor for wildland fire smoke. Atmos Meas Tech 11:1087–1097. doi: 10.5194/amt-11-1087-2018

Kelly KE, Whitaker J, Petty A, et al (2017) Ambient and laboratory evaluation of a low-cost particulate matter sensor. Environ Pollut 221:491–500. doi: 10.1016/J.ENVPOL.2016.12.039

Kim S-Y, Yi S-J, Eum YS, et al (2014) Ordinary kriging approach to predicting long-term particulate matter concentrations in seven major Korean cities. Environ Health Toxicol 29:e2014012. doi: 10.5620/eht.e2014012

Kim Y-M, Zhou Y, Gao Y, et al (2015) Spatially resolved estimation of ozone-related mortality in the United States under two representative concentration pathways (RCPs) and their uncertainty. Clim Change 128:71–84. doi: 10.1007/s10584-014-1290-1

Krall JR, Mulholland JA, Russell AG, et al (2017) Associations between Source-Specific Fine Particulate Matter and Emergency Department Visits for Respiratory Disease in Four U.S. Cities. Environ Health Perspect 125:97–103. doi: 10.1289/EHP271

Kukavskaya E a, Soja a J, Petkov a P, et al (2013) Fire emissions estimates in Siberia: evaluation of uncertainties in area burned, land cover, and fuel consumption. Can J For Res Can Rech For 43:493–506. doi: DOI 10.1139/cjfr-2012-0367

Kumar P, Morawska L, Martani C, et al (2015) The rise of low-cost sensing for managing air pollution in cities. Environ Int 75:199–205. doi: 10.1016/J.ENVINT.2014.11.019

Larkin NK, O'Neill SM, Solomon R, et al (2009) The BlueSky smoke modeling framework. Int J Wildl Fire 18:906. doi: 10.1071/WF07086

Lefohn AS, Knudsen HP, Logan JA, et al (1987) An Evaluation of the Kriging Method to Predict 7-h Seasonal Mean Ozone Concentrations for Estimating Crop Losses. JAPCA 37:595–602. doi: 10.1080/08940630.1987.10466247

Leopold A (1987) Game Management. Univ of Wisconsin Press

Li Z, Fraser R, Jin J, et al (2003) Evaluation of algorithms for fire detection and mapping across North America from satellite. J Geophys Res 108:4076. doi: 10.1029/2001JD001377

Li Z, Nadon S, Cihlar J (2000a) Satellite-based detection of Canadian boreal forest fires: Development and application of the algorithm. Int J Remote Sens 21:3057–3069. doi: 10.1080/01431160050144956

Li Z, Nadon S, Cihlar J, Stocks B (2000b) Satellite-based mapping of Canadian boreal forest fires: Evaluation and comparison of algorithms. Int J Remote Sens 21:3071–3082. doi: 10.1080/01431160050144965

Linares C, Carmona R, Tobías A, et al (2015) Influence of advections of particulate matter

from biomass combustion on specific-cause mortality in Madrid in the period 2004–2009. Environ Sci Pollut Res 22:7012–7019. doi: 10.1007/s11356-014-3916-2

Liu Y (2005) Enhancement of the 1988 northern U.S. drought due to wildfires. Geophys Res Lett 32:L10806. doi: 10.1029/2005GL022411

Liu Y (2014) A Regression Model for Smoke Plume Rise of Prescribed Fires Using Meteorological Conditions. J Appl Meteorol Climatol 53:1961–1975. doi: 10.1175/JAMC-D-13-0114.1

Liu Y, Koutrakis P, Kahn R, et al (2012) Estimating Fine Particulate Matter Component Concentrations and Size Distributions Using Satellite-Retrieved Fractional Aerosol Optical Depth: Part 2—A Case Study. J Air Waste Manage Assoc 57:1360–1369.

Liu Y, Sarnat JA, Kilaru V, et al (2005) Estimating Ground-Level PM 2.5 in the Eastern United States Using Satellite Remote Sensing. Environ Sci Technol 39:3269–3278. doi: 10.1021/es049352m

Liu Y, Stanturf J, Goodrick S (2010) Trends in global wildfire potential in a changing climate. For Ecol Manage 259:685–697. doi: 10.1016/J.FORECO.2009.09.002

Malm WC, Sisler JF, Huffman D, et al (1994) Spatial and seasonal trends in particle concentration and optical extinction in the United States. J Geophys Res 99:1347. doi: 10.1029/93JD02916

Marmur A, Unal A, Mulholland JA, Russell AG (2005) Optimization-Based Source Apportionment of PM 2.5 Incorporating Gas-to-Particle Ratios. Environ Sci Technol 39:3245–3254. doi: 10.1021/es0490121

Matte TD, Cohen A, Dimmick F, et al (2009) Summary of the workshop on methodologies for environmental public health tracking of air pollution effects. Air Qual Atmos Health 2:177–184. doi: 10.1007/s11869-009-0059-6

McGuinn LA, Ward-Caviness C, Neas LM, et al (2017) Fine particulate matter and cardiovascular disease: Comparison of assessment methods for long-term exposure. Environ Res 159:16–23. doi: 10.1016/j.envres.2017.07.041

McKenzie D, Raymond CL, Kellogg L-KB, et al (2007) Mapping fuels at multiple scales: landscape application of the Fuel Characteristic Classification SystemThis article is one of a selection of papers published in the Special Forum on the Fuel Characteristic Classification System. Can J For Res 37:2421–2437. doi: 10.1139/X07-056

McNider RT, Pour-Biazar A, Doty K, et al (2018) Examination of the Physical Atmosphere in the Great Lakes Region and its Potential Impact on Air Quality - Over-Water Stability and Satellite Assimilation. J Appl Meteorol Climatol JAMC–D–17–0355.1. doi: 10.1175/JAMC-D-17-0355.1

Mouillot F, Schultz MG, Yue C, et al (2014) Ten years of global burned area products from spaceborne remote sensing-A review: Analysis of user needs and recommendations for future developments. Int J Appl Earth Obs Geoinf 26:64–79. doi: 10.1016/j.jag.2013.05.014

Napelenok SL, Cohan DS, Hu Y, Russell AG (2006) Decoupled direct 3D sensitivity analysis for particulate matter (DDM-3D/PM). Atmos Environ 40:6112–6121. doi: 10.1016/J.ATMOSENV.2006.05.039

Natarajan M, Pierce RB, Schaack TK, et al (2012) Radiative forcing due to enhancements in tropospheric ozone and carbonaceous aerosols caused by Asian fires during spring 2008. J Geophys Res Atmos 117:n/a–n/a. doi: 10.1029/2011JD016584

Odman M, Huang R, Pophale A, et al (2018) Forecasting the Impacts of Prescribed Fires for Dynamic Air Quality Management. Atmosphere (Basel) 9:220. doi: 10.3390/atmos9060220

Odman MT, Chang ME, Hu Y, et al (2017) Final Report: Dynamic Management of Prescribed Burning for Better Air Quality.

Odman MT, Qin M, Hu Y, et al (2019) 2017 Projections and Interstate Transport of Ozone in the Southeastern United States.

Pachon JE, Balachandran S, Hu Y, et al (2012) Development of outcome-based, multipollutant mobile source indicators. J Air Waste Manage Assoc 62:431–442. doi: 10.1080/10473289.2012.656218

Park Williams A, Cook BI, Smerdon JE, et al (2017) The 2016 Southeastern U.S. Drought: An Extreme Departure From Centennial Wetting and Cooling. J Geophys Res Atmos 122:10,888–10,905. doi: 10.1002/2017JD027523

Pleim J, Gilliam R, Appel W, Ran L (2016) Recent Advances in Modeling of the Atmospheric Boundary Layer and Land Surface in the Coupled WRF-CMAQ Model. Springer International Publishing, pp 391–396

Pope CA, Ezzati M, Dockery DW (2009) Fine-Particulate Air Pollution and Life Expectancy in the United States. N Engl J Med 360:376–386. doi:

10.1056/NEJMsa0805646

Prins EM, Menzel WP (1992) Geostationary satellite detection of bio mass burning in South America. Int J Remote Sens 13:2783–2799. doi: 10.1080/01431169208904081

Prins EM, Menzel WP (1994) Trends in South American biomass burning detected with the GOES visible infrared spin scan radiometer atmospheric sounder from 1983 to 1991. J Geophys Res 99:16719. doi: 10.1029/94JD01208

Qin M, Wang X, Hu Y, et al (2015) Formation of particulate sulfate and nitrate over the Pearl River Delta in the fall: Diagnostic analysis using the Community Multiscale Air Quality model. Atmos Environ 112:81–89. doi: 10.1016/j.atmosenv.2015.04.027

Rai AC, Kumar P, Pilla F, et al (2017) End-user perspective of low-cost sensors for outdoor air pollution monitoring. Sci Total Environ 607-608:691–705. doi: 10.1016/J.SCITOTENV.2017.06.266

Randerson JT, Chen Y, Van Der Werf GR, et al (2012) Global burned area and biomass burning emissions from small fires. J Geophys Res Biogeosciences. doi: 10.1029/2012JG002128

Randerson JT, Liu H, Flanner MG, et al (2006) The Impact of Boreal Forest Fire on Climate Warming.

Rappold AG, Reyes J, Pouliot G, et al (2017) Community Vulnerability to Health Impacts of Wildland Fire Smoke Exposure. Environ Sci Technol 51:6674–6682. doi: 10.1021/acs.est.6b06200

Rappold AG, Stone SL, Cascio WE, et al (2011) Peat Bog Wildfire Smoke Exposure in Rural North Carolina Is Associated with Cardiopulmonary Emergency Department Visits Assessed through Syndromic Surveillance. Environ Health Perspect 119:1415–1420. doi: 10.1289/ehp.1003206

Reid CE, Brauer M, Johnston FH, et al (2016) Critical Review of Health Impacts of Wildfire Smoke Exposure. Environ Health Perspect 124:1334–1343. doi: 10.1289/ehp.1409277

Reisen F, Meyer CP (Mick), McCaw L, et al (2011) Impact of smoke from biomass burning on air quality in rural communities in southern Australia. Atmos Environ 45:3944–3953. doi: 10.1016/J.ATMOSENV.2011.04.060

Rolph GD, Draxler RR, Stein AF, et al (2009) Description and Verification of the NOAA

Smoke Forecasting System: The 2007 Fire Season. Weather Forecast 24:361–378. doi: 10.1175/2008WAF2222165.1

Roy DP, Boschetti L, Justice CO, Ju J (2008) The collection 5 MODIS burned area product — Global evaluation by comparison with the MODIS active fire product. Remote Sens Environ 112:3690–3707. doi: 10.1016/j.rse.2008.05.013

Sacks JD, Lloyd JM, Zhu Y, et al (2018) The Environmental Benefits Mapping and Analysis Program – Community Edition (BenMAP–CE): A tool to estimate the health and economic benefits of reducing air pollution. Environ Model Softw 104:118–129. doi: 10.1016/J.ENVSOFT.2018.02.009

Sampson PD, Richards M, Szpiro AA, et al (2013) A regionalized national universal kriging model using Partial Least Squares regression for estimating annual PM2.5 concentrations in epidemiology. Atmos Environ (1994) 75:383–392. doi: 10.1016/j.atmosenv.2013.04.015

Sarnat JA, Marmur A, Klein M, et al (2008) Fine Particle Sources and Cardiorespiratory Morbidity: An Application of Chemical Mass Balance and Factor Analytical Source-Apportionment Methods. Environ Health Perspect 116:459–466. doi: 10.1289/ehp.10873

Sarnat SE, Coull BA, Schwartz J, et al (2005) Factors Affecting the Association between Ambient Concentrations and Personal Exposures to Particles and Gases. Environ Health Perspect 114:649–654. doi: 10.1289/ehp.8422

Schroeder W, Csiszar I, Giglio L, Schmidt CC (2010) On the use of fire radiative power, area, and temperature estimates to characterize biomass burning via moderate to coarse spatial resolution remote sensing data in the Brazilian Amazon. J Geophys Res 115:D21121. doi: 10.1029/2009JD013769

Schultz MG (2002) On the use of ATSR fire count data to estimate the seasonal and interannual variability of vegetation fire emissions. Atmos Chem Phys Atmos Chem Phys 2:387–395.

Seiler W, Crutzen PJ (1980) Estimates of gross and net fluxes of carbon between the biosphere and the atmosphere from biomass burning. Clim Change 2:207–247. doi: 10.1007/BF00137988

Simon M, Plummer S, Fierens F, et al (2004) Burnt area detection at global scale using ATSR-2: The GLOBSCAR products and their qualification. J Geophys Res

109:D14S02. doi: 10.1029/2003JD003622

Snyder EG, Watkins TH, Solomon PA, et al (2013) The Changing Paradigm of Air Pollution Monitoring. Environ Sci Technol 47:11369–11377. doi: 10.1021/es4022602

Sokolik IN, Curry J, Radionov V (2010) Interactions of Arctic Aerosols with Land-Cover and Land-Use Changes in Northern Eurasia and their Role in the Arctic Climate System. In: Eurasian Arctic Land Cover and Land Use in a Changing Climate. Springer Netherlands, Dordrecht, pp 237–268

Solomon PA, Costantini M, Grahame TJ, et al (2012) Air pollution and health: bridging the gap from sources to health outcomes: conference summary. Air Qual Atmos Heal 5:9–62. doi: 10.1007/S11869-011-0161-4

Stein AF, Draxler RR, Rolph GD, et al (2015) NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. Bull Am Meteorol Soc 96:2059–2077. doi: 10.1175/BAMS-D-14-00110.1

Tang W, Cohan DS, Morris GA, et al (2011) Influence of vertical mixing uncertainties on ozone simulation in CMAQ. Atmos Environ 45:2898–2909. doi: 10.1016/j.atmosenv.2011.01.057

Tansey K, Beston J, Hoscilo A, et al (2008a) Relationship between MODIS fire hot spot count and burned area in a degraded tropical peat swamp forest in Central Kalimantan, Indonesia. J Geophys Res 113:D23112. doi: 10.1029/2008JD010717

Tansey K, Grégoire J, Stroppiana D, et al (2004) Vegetation burning in the year 2000: Global burned area estimates from SPOT VEGETATION data. J Geophys Res 109:D14S03. doi: 10.1029/2003JD003598

Tansey K, Grégoire J-M, Defourny P, et al (2008b) A new, global, multi-annual (2000–2007) burnt area product at 1 km resolution. Geophys Res Lett 35:L01401. doi: 10.1029/2007GL031567

US EPA (2014) 2014 National Emissions Inventory (NEI) Documentation.

US EPA BenMAP Community Edition. https://www.epa.gov/benmap/benmap-community-edition. Accessed 15 Feb 2019

Van Donkelaar A, Martin R V., Park RJ, et al (2007) Model evidence for a significant source of secondary organic aerosol from isoprene. Atmos Environ 41:1267–1274.

doi: 10.1016/j.atmosenv.2006.09.051

Wade KS, Mulholland JA, Marmur A, et al (2006) Effects of Instrument Precision and Spatial Variability on the Assessment of the Temporal Variation of Ambient Air Pollution in Atlanta, Georgia. J Air Waste Manage Assoc 56:876–888. doi: 10.1080/10473289.2006.10464499

Wagstrom KM, Pandis SN, Yarwood G, et al (2008) Development and application of a computationally efficient particulate matter apportionment algorithm in a three-dimensional chemical transport model. Atmos Environ 42:5650–5659. doi: 10.1016/J.ATMOSENV.2008.03.012

Woody MC, Baker KR, Hayes PL, et al (2016) Understanding sources of organic aerosol during CalNex-2010 using the CMAQ-VBS. Atmos Chem Phys 16:4081–4100. doi: 10.5194/acp-16-4081-2016

Wyat Appel K, Bhave P V., Gilliland AB, et al (2008) Evaluation of the community multiscale air quality (CMAQ) model version 4.5: Sensitivities impacting model performance; Part II—particulate matter. Atmos Environ 42:6057–6066. doi: 10.1016/j.atmosenv.2008.03.036

Xiao Q, Liu Y, Mulholland JA, et al (2016) Pediatric emergency department visits and ambient Air pollution in the U.S. State of Georgia: a case-crossover study. Environ Heal 15:115. doi: 10.1186/s12940-016-0196-y

Xiao X, Cohan DS, Byun DW, Ngan F (2010) Highly nonlinear ozone formation in the Houston region and implications for emission controls. J Geophys Res 115:D23309. doi: 10.1029/2010JD014435

Yao J, Eyamie J, Henderson SB (2016) Evaluation of a spatially resolved forest fire smoke model for population-based epidemiologic exposure assessment. J Expo Sci Environ Epidemiol 26:233–240. doi: 10.1038/jes.2014.67

Yu S, Mathur R, Pleim J, et al (2012) Comparative evaluation of the impact of WRF/NMM and WRF/ARW meteorology on CMAQ simulations for $PM_{2.5}$ and its related precursors during the 2006 TexAQS/GoMACCS study. Atmos Chem Phys 12:4091–4106. doi: 10.5194/acp-12-4091-2012

Zeng T, Liu Z, Wang Y (2016) Large fire emissions in summer over the southeastern US: Satellite measurements and modeling analysis. Atmos Environ 127:213–220. doi: 10.1016/j.atmosenv.2015.12.025

Zhang X, Kondragunta S (2008) Temporal and spatial variability in biomass burned areas across the USA derived from the GOES fire product. Remote Sens Environ 112:2886–2897. doi: 10.1016/j.rse.2008.02.006

Zhang X, Kondragunta S, Quayle B (2011) Estimation of Biomass Burned Areas Using Multiple-Satellite-Observed Active Fires. IEEE Trans Geosci Remote Sens 49:4469–4482. doi: 10.1109/TGRS.2011.2149535

Zhang Y, Huang J-P, Henze DK, Seinfeld JH (2007) Role of isoprene in secondary organic aerosol formation on a regional scale. J Geophys Res 112:D20207. doi: 10.1029/2007JD008675

Zheng T, Bergin MH, Johnson KK, et al (2018) Field evaluation of low-cost particulate matter sensors in high and low concentration environments. Atmos Meas Tech Discuss 1–40. doi: 10.5194/amt-2018-111

Zhu C, Kobayashi H, Kanaya Y, Saito M (2017) Sizea-dependent validation of MODIS MCD64A1 burned area over six vegetation types in boreal Eurasia: Large underestimation in croplands. Sci Rep 7:4181. doi: 10.1038/s41598-017-03739-0